

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
21 October 2004 (21.10.2004)

PCT

(10) International Publication Number  
**WO 2004/089983 A2**

- (51) International Patent Classification<sup>7</sup>: **C07K 14/65**, C12N 15/85, 5/10, C07K 16/10, C12Q 1/70, G01N 33/569, A61K 39/215, 38/21
- (74) Agent: WINCKELS, J., H., F.; Nieuwe Parklaan 97, NL-2587 BN Den Haag (NL).
- (21) International Application Number:  
PCT/NL2004/000229
- (22) International Filing Date: 8 April 2004 (08.04.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
03076037.5 8 April 2003 (08.04.2003) EP  
03076110.0 14 April 2003 (14.04.2003) EP  
03077307.1 18 July 2003 (18.07.2003) EP
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (71) Applicant (for all designated States except US): **VI-RONOVATIVE B.V.** [NL/NL]; Burgemeester Oudlaan 50, NL-3062 PA Rotterdam (NL).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **HAAGMANS, Bartholomeus, Leonardus** [NL/NL]; Doude van Troostwijkstraat 3, NL-1391 EP Abcoude (NL). **KUIKEN, Thijs** [NL/NL]; Middenmolenlaan 149, NL-2807 EV Gouda (NL). **FOUCHIER, Ronaldus, Adrianus, Maria** [NL/NL]; Essenburgsingel 44a, NL-3021 AR Rotterdam (NL). **OSTERHAUS, Albertus, Dominicus, Marcellinus, Erasmus** [NL/NL]; Dr. Breveestraat 16, NL-3981 CH Bunnik (NL).
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

- without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/089983 A2

(54) Title: SARS

(57) Abstract: The invention relates to the field of virology. The invention provides an isolated essentially mammalian positive-sense single stranded RNA virus (SARS) within the group of coronaviruses and components thereof.

Title: SARS

5

The invention relates to the field of virology, more in particular to a new coronavirus. In particular sequences encoding (parts of) viral proteins are provided. Further, the invention relates to diagnostic means and methods, prophylactic means and methods and therapeutic means and methods to be employed in the diagnosis, prevention and/or treatment of disease, in particular of respiratory disease (atypical pneumonia), in particular of mammals, more in particular in humans. In another embodiment the invention relates to the use of interferon, preferably pegylated interferon for the prophylactic or therapeutic treatment of animals, preferably vertebrates, more preferably birds or mammals, especially human, apes or rodents, infected with a coronavirus, more specifically an animal, preferably human infected with a SARS associated coronavirus (SARS-CoV).

Recently, a new virus has caused a global health risk because of its pathogenic effects in man combined with a relatively easy droplet transmission. The virus first was seen in the Chinese province Guangdong, was spread to Hong Kong in February 2003, and within two months it has been able to spread to several countries all over the world where it has caused 78 deaths out of 2300 people infected (New Scientist Online News 13:25 02 April 2003). The virus has been named SARS (Severe Acute Respiratory Syndrome) virus and causes a respiratory illness (atypical pneumonia) in man. This illness usually begins with a fever, sometimes associated with chills or other symptoms, including headache, rash, diarrhea, a general feeling of discomfort (malaise) and body aches. Some people also experience mild respiratory syndromes at the outset.

After 2 to 7 days, SARS patients may develop a dry, nonproductive cough that might be accompanied or progress to the point where insufficient oxygen is getting to the blood, visible as shortness of breath. In 10% to 20% of the cases, patients will require mechanical ventilation, and eventually the disease can lead to the death of the patient. Hospital personnel, children, elderly and people having an underlying condition such as diabetes or heart disease, or a weakened immune system, form the

highest risk group. Co-infection with other pathogens seems to occur frequently, especially with opportunistic pathogenic microorganisms such as human metapneumovirus (hMPV), Chlamydia, etcetera.

The incubation time for the virus is typically 2-7 days and the disease is transmitted by people sick with SARS coughing or sneezing droplets in the air.

As for yet it is not known if there is a cure for the disease. Several antiviral therapies have been applied, but with various results.

Also, for being able to prevent spread of the disease, it is of great importance to be able to recognise the disease in an early stage. Only then sufficient measures can be taken to isolate patients and initiate quarantine precautions. At this moment there is not yet a diagnostic tool in place.

Thus, there is great need in developing diagnostic tools and therapies for this disease.

The invention provides the nucleotide sequence of an isolated essentially mammalian positive-sense single stranded RNA virus belonging to the Coronaviruses, which is the causative factor for SARS. From a phylogenetic analysis of the sequences of the virus (Fig. 1) it appears that the virus is an intermediate between the group formed by TGEV (transmissible gastroenteritis virus), PEDV (porcine epidemic diarrhea virus) and 229E (human coronavirus 229E) at one side, the group formed by BoCo (bovine coronavirus) and MHV (murine hepatitis virus) at an other side, and the AIBV (avian infectious bronchitis virus) on yet another side. In general, bovine coronavirus seems to be the closest relative (at least for the viral replicase protein).

Although phylogenetic analyses provide a convenient method of identifying a virus as a SARS virus several other possibly more straightforward albeit somewhat more coarse methods for identifying said virus or viral proteins or nucleic acids from said virus are herein also provided. As a rule of thumb a SARS virus can be identified by the percentages of homology of the virus, proteins or nucleic acids to be identified in comparison with viral proteins or nucleic acids identified herein by sequence. It is generally known that virus species, especially RNA virus species, often constitute a quasi species wherein a cluster of said viruses displays heterogeneity among its members. Thus it is expected that each isolate may have a somewhat different percentage relationship with the sequences of the isolate as provided herein.

When one wishes to compare a virus isolate with the sequences as listed in figure 2, the invention provides an isolated essentially mammalian positive-sense single stranded RNA virus (SARS) belonging to the Coronaviruses and identifiable as phylogenetically corresponding thereto by determining a nucleic acid sequence of said virus and determining that said nucleic acid sequence has a percentage nucleic acid identity to the sequences as listed higher than the percentages identified herein for the nucleic acids as identified herein below in comparison with BoCo, AIPV and PEDV. Likewise, an isolated essentially mammalian positive-sense single stranded RNA virus (SARS) belonging to the Coronaviruses and identifiable as phylogenetically corresponding thereto by determining an amino acid sequence of said virus and determining that said amino acid sequence has a percentage amino acid homology to the sequences as listed which is essentially higher than the percentages provided herein in comparison with BoCo, AIPV and PEDV.

With the provision of the sequence information of this SARS virus, the invention provides diagnostic means and methods, prophylactic means and methods and therapeutic means and methods to be employed in the diagnosis, prevention and/or treatment of disease, in particular of respiratory disease (atypical pneumonia), in particular of mammals, more in particular in humans. In virology, it is most advisory that diagnosis, prophylaxis and/or treatment of a specific viral infection is performed with reagents that are most specific for said specific virus causing said infection. In this case this means that it is preferred that said diagnosis, prophylaxis and/or treatment of a SARS virus infection is performed with reagents that are most specific for SARS virus. This by no means however excludes the possibility that less specific, but sufficiently cross-reactive reagents are used instead, for example because they are more easily available and sufficiently address the task at hand. The invention for example provides a method for virologically diagnosing a SARS infection of an animal, in particular of a mammal, more in particular of a human being, comprising determining in a sample of said animal the presence of a viral isolate or component thereof by reacting said sample with a SARS specific nucleic acid or antibody according to the invention, and a method for serologically diagnosing a SARS infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a SARS virus or component



thereof by reacting said sample with a SARS virus-specific proteinaceous molecule or fragment thereof or an antigen according to the invention.

The invention also provides a diagnostic kit for diagnosing a SARS infection comprising a SARS virus, a SARS virus-specific nucleic acid, proteinaceous molecule  
5 or fragment thereof, antigen and/or an antibody according to the invention, and preferably a means for detecting said SARS virus, SARS virus-specific nucleic acid, proteinaceous molecule or fragment thereof, antigen and/or an antibody, said means for example comprising an excitable group such as a fluorophore or enzymatic detection system used in the art (examples of suitable diagnostic kit format comprise  
10 IF, ELISA, neutralization assay, RT-PCR assay). To determine whether an as yet unidentified virus component or synthetic analogue thereof such as nucleic acid, proteinaceous molecule or fragment thereof can be identified as SARS-virus-specific, it suffices to analyse the nucleic acid or amino acid sequence of said component, for example for a stretch of said nucleic acid or amino acid, preferably of at least 10,  
15 more preferably at least 25, more preferably at least 40 nucleotides or amino acids (respectively), by sequence homology comparison with the provided SARS viral sequences and with known non-SARS viral sequences (BoCo is preferably used) using for example phylogenetic analyses as provided herein. Depending on the degree of relationship with said SARS or non-SARS viral sequences, the component or  
20 synthetic analogue can be identified.

The invention thus provides the nucleotide sequence of a novel etiological agent, an isolated essentially mammalian positive-sense single stranded RNA virus (herein also called SARS virus) belonging to the Coronaviridae family, and SARS  
25 virus-specific components or synthetic analogues thereof. Coronaviruses were first isolated from chickens in 1937, while the first human coronavirus was propagated *in vitro* by Tyrell and Bonoe in 1965. There are now about 13 species in this family, which infect cattle, pigs, rodents, cats, dogs, birds and man. Coronavirus particles are irregularly shaped, about 60-220 nm in diameter, with an outer envelope bearing  
30 distinctive, 'club-shaped' peplomers (about 20 nm long and 10 nm wide at the distal end). This 'crown-like' appearance give the family its name. The envelope carries two glycoproteins: S, the spike glycoprotein which is involved in cell fusion and is a major antigen, and M, the membrane glycoprotein, which is involved in budding and envelope formation. The genome is associated with a basic phosphoprotein,

designated N. The genome of coronaviruses, a single stranded positive-sense RNA strand, is typically 27-31 Kb long and contains a 5' methylated cap and a 3' poly-A tail, by which it can directly function as an mRNA in the infected cell. Initially the 5' ORF 1 (about 20 Kb) is translated to produce a viral polymerase, which then produces  
5 a full length negative sense strand. This is used as a template to produce mRNA as a 'nested set' of transcripts, all with identical 5' non-translated leader sequence of 72 nucleotides and coincident 3' polyadenylated ends. Each mRNA thus produced is monocistronic, the genes at the 5' end being translated from the longest mRNA and so on. These unusual cytoplasmic structures are produced not by splicing, but by the  
10 polymerase during transcription. Between each of the genes there is a repeated intergenic sequence – AACUAAAC – which interacts with the transcriptase plus cellular factors to splice the leader sequence onto the start of each ORF. In some coronaviruses there are about 8 ORFs, coding for the proteins mentioned above, but also for a hemagglutinin esterase (HE), and several other non-structural proteins.  
15 Newly isolated viruses are phylogenetically corresponding to and thus taxonomically corresponding to SARS virus when comprising a gene order and/or amino acid sequence and/or nucleotide sequence sufficiently similar to our prototypic SARS virus. The highest amino acid sequence homology, between SARS virus and any of the known other viruses of the same family to date (BoCo or Mouse Hepatitis Virus) is  
20 for parts of the polymerase protein 18-61% (the % homology, and the virus to which the homology is depend on the region of the polymerase that is examined), as can be deduced when comparing the sequences given in figure 2 with sequences of other viruses, in particular of BoCo and Mouse Hepatitis Virus. Individual proteins or whole virus isolates with, respectively, higher homology than these mentioned  
25 maximum values are considered phylogenetically corresponding and thus taxonomically corresponding to SARS virus, and generally will be encoded by a nucleic acid sequence structurally corresponding with a sequence as shown in figure 2. Herewith the invention provides a virus phylogenetically corresponding to the isolated virus of which the sequences are depicted in figure 2.  
30 It should be noted that, similar to other viruses, a certain degree of variation can be expected to be found between SARS-viruses isolated from different sources. Also, the viral sequence of the SARS virus or an an isolated SARS virus gene as provided herein for example shows less than 95%, preferably less than 90%, more preferably less than 80%, more preferably less than 70% and most preferably less

than 65% nucleotide sequence homology or less than 95%, preferably less than 90%, more preferably less than 80%, more preferably less than 70% and most preferably less than 65% amino acid sequence homology with the respective nucleotide or amino acid sequence of the bovine coronavirus or the murine hepatitis virus as for example  
5 can be found in Genbank (for example in accession number NC\_002306 (BoCo) or NC\_002645 (MHV)).

Sequence divergence of SARS strains around the world may be somewhat higher, in analogy with other coronaviruses.

A fair number of virus isolates have been isolated during the priority year of the  
10 present application, and it has been found that these viruses share the homology indicated above. The sequences of these viruses can be found in GenBank accession no. AY274119 (see fig. 10) or AY278741 or AY338175 or AY338174 or AY322199 or AY 322198 or AY322197 or AH013000 or AY322208 or AY322207 AY 322206 or AY322205 or AH012999 and and/or sequences depicted in  
15 <http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Undef&id=227859&lvl=3&keep=1&srchmode=1&unlock>. Herewith the invention encompasses a virus phylogenetically corresponding to the isolated virus of which the sequences are depicted in figure 2 and/or for example the GenBank accession no. AY274119 or AY278741 or AY338175 or AY338174 or AY322199 or AY 322198 or AY322197 or  
20 AH013000 or AY322208 or AY322207 AY 322206 or AY322205 or AH012999 and and/or sequences depicted in  
<http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Undef&id=227859&lvl=3&keep=1&srchmode=1&unlock>.

25 The term "nucleotide sequence homology" as used herein denotes the presence of homology between two (poly)nucleotides. Polynucleotides have "homologous" sequences if the sequence of nucleotides in the two sequences is the same when aligned for maximum correspondence. Sequence comparison between two or more polynucleotides is generally performed by comparing portions of the two sequences  
30 over a comparison window to identify and compare local regions of sequence similarity. The comparison window is generally from about 20 to 200 contiguous nucleotides. The "percentage of sequence homology" for polynucleotides, such as 50, 60, 70, 80, 90, 95, 98, 99 or 100 percent sequence homology may be determined by comparing two optimally aligned sequences over a comparison window, wherein the

portion of the polynucleotide sequence in the comparison window may include additions or deletions (i.e. gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by: (a) determining the number of positions at which the identical nucleic acid base occurs in both sequences to yield the number of matched positions; (b) dividing the number of matched positions by the total number of positions in the window of comparison; and (c) multiplying the result by 100 to yield the percentage of sequence homology. Optimal alignment of sequences for comparison may be conducted by computerized implementations of known algorithms, or by inspection. Readily available sequence comparison and multiple sequence alignment algorithms are, respectively, the Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. 1990. J. Mol. Biol. 215:403; Altschul, S.F. et al. 1997. Nucleic Acid Res. 25:3389-3402) and ClustalW programs both available on the internet. Other suitable programs include GAP, BESTFIT and FASTA in the Wisconsin Genetics Software Package (Genetics Computer Group (GCG), Madison, WI, USA).

As used herein, "substantially complementary" means that two nucleic acid sequences have at least about 65%, preferably about 70%, more preferably about 80%, even more preferably 90%, and most preferably about 98%, sequence complementarity to each other. This means that the primers and probes must exhibit sufficient complementarity to their template and target nucleic acid, respectively, to hybridise under stringent conditions. Therefore, the primer sequences as disclosed in this specification need not reflect the exact sequence of the binding region on the template and degenerate primers can be used. A substantially complementary primer sequence is one that has sufficient sequence complementarity to the amplification template to result in primer binding and second-strand synthesis.

The term "hybrid" refers to a double-stranded nucleic acid molecule, or duplex, formed by hydrogen bonding between complementary nucleotides. The terms "hybridise" or "anneal" refer to the process by which single strands of nucleic acid sequences form double-helical segments through hydrogen bonding between complementary nucleotides.

The term "oligonucleotide" refers to a short sequence of nucleotide monomers (usually 6 to 100 nucleotides) joined by phosphorous linkages (e.g., phosphodiester, alkyl and aryl-phosphate, phosphorothioate), or non-phosphorous linkages (e.g., peptide, sulfamate and others). An oligonucleotide may contain modified nucleotides having

modified bases (e.g., 5-methyl cytosine) and modified sugar groups (e.g., 2'-O-methyl ribosyl, 2'-O-methoxyethyl ribosyl, 2'-fluoro ribosyl, 2'-amino ribosyl, and the like).

Oligonucleotides may be naturally-occurring or synthetic molecules of double- and single-stranded DNA and double- and single-stranded RNA with circular, branched or linear shapes and optionally including domains capable of forming stable secondary structures (e.g., stem-and-loop and loop-stem-loop structures).

The term "primer" as used herein refers to an oligonucleotide which is capable of annealing to the amplification target allowing a DNA polymerase to attach thereby serving as a point of initiation of DNA synthesis when placed under conditions in

which synthesis of primer extension product which is complementary to a nucleic acid strand is induced, i.e., in the presence of nucleotides and an agent for polymerization such as DNA polymerase and at a suitable temperature and pH. The (amplification) primer is preferably single stranded for maximum efficiency in amplification.

Preferably, the primer is an oligodeoxy ribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the agent for polymerization. The exact lengths of the primers will depend on many factors, including temperature and source of primer. A "pair of bi-directional primers" as used herein refers to one forward and one reverse primer as commonly used in the art of DNA amplification such as in PCR amplification.

The term "probe" refers to a single-stranded oligonucleotide sequence that will recognize and form a hydrogen-bonded duplex with a complementary sequence in a target nucleic acid sequence analyte or its cDNA derivative.

The terms "stringency" or "stringent hybridization conditions" refer to hybridization conditions that affect the stability of hybrids, e.g., temperature, salt concentration, pH, formamide concentration and the like. These conditions are empirically optimised to maximize specific binding and minimize non-specific binding of primer or probe to its target nucleic acid sequence. The terms as used include reference to conditions under which a probe or primer will hybridise to its target sequence, to a detectably greater degree than other sequences (e.g. at least 2-fold over background). Stringent

conditions are sequence dependent and will be different in different circumstances. Longer sequences hybridise specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of a complementary target

sequence hybridises to a perfectly matched probe or primer. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M Na<sup>+</sup> ion, typically about 0.01 to 1.0 M Na<sup>+</sup> ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes or primers (e.g. 10 to 50 nucleotides) and at least about 60°C for long probes or primers (e.g. greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringent conditions or "conditions of reduced stringency" include hybridization with a buffer solution of 30% formamide, 1 M NaCl, 1% SDS at 37°C and a wash in 2x SSC at 40°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1x SSC at 60°C. Hybridization procedures are well known in the art and are described in e.g. Ausubel et al, Current Protocols in Molecular Biology, John Wiley & Sons Inc., 1994.

The term "antibody" includes reference to antigen binding forms of antibodies (e. g., Fab, F (ab) 2). The term "antibody" frequently refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen). However, while various antibody fragments can be defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized de novo either chemically or by utilizing recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments such as single chain Fv, chimeric antibodies (i. e., comprising constant and variable regions from different species), humanized antibodies (i. e., comprising a complementarity determining region (CDR) from a non-human source) and heteroconjugate antibodies (e. g., bispecific antibodies).

"Interferon" is a term generically comprehending a group of vertebrate glycoproteins and proteins which are known to have various biological activities, such as antiviral, antiproliferative, and immunomodulatory activity at least in the species of animal from which such substances are derived. Interferon refers to a class of small protein and glycoprotein cytokines (15–28 kD) produced by T cells, fibroblasts, and other cells in response to viral infection and other biological and synthetic stimuli. Interferons bind to specific receptors on cell membranes; their effects include inducing enzymes, suppressing cell proliferation, inhibiting viral proliferation, enhancing the phagocytic activity of macrophages, and augmenting the cytotoxic activity of T lymphocytes.

Interferons are divided into five major classes (alpha, beta, gamma, tau, and omega) and several subclasses (indicated by Arabic numerals and letters) on the basis of physicochemical properties, cells of origin, mode of induction, and antibody reactions.

5           In short, the invention provides an isolated essentially mammalian positive-sense single stranded RNA virus (SARS) belonging to the Coronaviruses and identifiable as phylogenetically corresponding thereto by determining a nucleic acid sequence of a suitable fragment of the genome of said virus and testing it in phylogenetic tree analyses wherein maximum likelihood trees are generated using  
10 100 bootstraps and 3 jumbles and finding it to be more closely phylogenetically corresponding to a virus isolate having the sequences as depicted in figure 2 than it is corresponding to a virus isolate of BoCo (bovine coronavirus, e.g. acc. no. NC\_002306 in Genbank), MHV (murine hepatitis virus, e.g. acc. no. NC\_002645), AIBV (avian infectious bronchitis virus, e.g. acc. no. NC\_001451), PEDV (porcine epidemic  
15 diarrhea virus), TGEV (transmissible gastroenteritis virus, e.g. acc. no. NC\_003436) or 229E (human coronavirus 229E, e.g. acc. no. NC\_003045). All the viral sequences with the GenBank accession numbers mentioned above are believed to be phylogenetically coresponding viruses to the virus of which the sequences are depicted in fig. 2.

20           Suitable nucleic acid genome fragments each useful for such phylogenetic tree analyses are for example any of the RAP-PCR fragments EMC-1 to -14 and RDG-1 as disclosed in figure 2, leading to the phylogenetic tree analysis as disclosed in figure 1.

A suitable open reading frame (ORF) comprises the ORF encoding the viral  
25 polymerase (ORF 1a). When an overall amino acid identity of at least 60%, preferably of at least 70%, more preferably of at least 80%, more preferably of at least 90%, most preferably of at least 95% of the analysed polymerase with the polymerase having a sequence comprising the amino acid fragments EMC-1, EMC-2, EMC-3, EMC-4, EMC-5, EMC-13 and/or EMC-14 of figure 2 is found, the analysed virus isolate comprises a  
30 SARS virus isolate according to the invention.

Another suitable open reading frame (ORF) useful in phylogenetic analyses comprises the ORF encoding the N protein. When an overall amino acid identity of at least 60%, more preferably of at least 70%, more preferably of at least 80%, more preferably of at least 90%, most preferably of at least 95% of the analysed N-protein

with the N-protein encoded by a sequence comprising the sequence EMC-8 of figure 2 is found, the analysed virus isolate comprises a SARS isolate according to the invention.

Another suitable open reading frame (ORF) useful in phylogenetic analyses  
5 comprises the ORF encoding the spike protein S. When an overall amino acid identity of at least 60%, more preferably of at least 70%, more preferably of at least 80%, more preferably of at least 90%, most preferably of at least 95% of the analysed S-protein encoded by a sequence comprising the sequence of translation 2 of EMC7 and translation 1 of the RDG 1 sequence of the S-protein as depicted in figure 2 is found,  
10 the analysed virus isolate comprises a SARS virus isolate according to the invention. The S ORF of the SARS virus seems to be located adjacent to the ORF 1ab (coding for the viral polymerase), which would discriminate SARS viruses from the bovine coronavirus and the murine hepatitis virus, which have a so-called 2a gene and an HE-gene between the S protein and the viral polymerase.

15

The invention provides among others an isolated or recombinant nucleic acid or virus-specific functional fragment thereof obtainable from a virus according to the invention. The isolated or recombinant nucleic acids comprises the sequences as given in figure 2 or sequences of homologues which are able to hybridise with those under  
20 stringent conditions. In particular, the invention provides primers and/or probes suitable for identifying a SARS virus nucleic acid.

Furthermore, the invention provides a vector comprising a nucleic acid according to the invention. To begin with, vectors such as plasmid vectors containing (parts of) the genome of SARS virus, virus vectors containing (parts of) the genome of SARS (for  
25 example, but not limited thereto, vaccinia virus, retroviruses, baculovirus), or SARS virus containing (parts of) the genome of other virus or other pathogens are provided.

Also, the invention provides a host cell comprising a nucleic acid or a vector according to the invention. Plasmid or viral vectors containing the polymerase components of  
30 SARS virus are generated in prokaryotic cells for the expression of the components in relevant cell types (bacteria, insect cells, eukaryotic cells). Plasmid or viral vectors containing full-length or partial copies of the SARS virus genome will be generated in prokaryotic cells for the expression of viral nucleic acids *in-vitro* or *in-vivo*. The latter vectors may contain other viral sequences for the generation of chimeric viruses or



chimeric virus proteins, may lack parts of the viral genome for the generation of replication defective virus, and may contain mutations, deletions or insertions for the generation of attenuated viruses.

Infectious copies of SARS virus (being wild type, attenuated, replication-defective or  
5 chimeric) can be produced upon co-expression of the polymerase components according to the state-of-the-art technologies described above.

In addition, eukaryotic cells, transiently or stably expressing one or more full-length or partial SARS virus proteins can be used. Such cells can be made by transfection (proteins or nucleic acid vectors), infection (viral vectors) or transduction (viral  
10 vectors) and may be useful for complementation of mentioned wild type, attenuated, replication-defective or chimeric viruses.

A chimeric virus may be of particular use for the generation of recombinant vaccines protecting against two or more viruses. For example, it can be envisaged that a SARS virus vector expressing one or more proteins of a human  
15 metapneumovirus or a human metapneumovirus vector expressing one or more proteins of SARS virus will protect individuals vaccinated with such vector against both virus infections. Such a specific chimeric virus is particularly useful in the invention because it is suspected that co-infection of, for instance, human metapneumovirus frequently occurs in SARS virus infected patients. Attenuated and  
20 replication-defective viruses may be of use for vaccination purposes with live vaccines as has been suggested for other viruses. Recently, Subbarao, K. *et al.*, *J. Virol.* 78(7), 3572-3577, 2004) demonstrated that mice are protected as a result from a previous immunisation with whole viruses.

In a preferred embodiment, the invention provides a proteinaceous molecule or  
25 coronavirus-specific viral protein or functional fragment thereof encoded by a nucleic acid according to the invention. Useful proteinaceous molecules are for example derived from any of the genes or genomic fragments derivable from a virus according to the invention. Such molecules, or antigenic fragments thereof, as provided herein, are for example useful in diagnostic methods or kits and in pharmaceutical  
30 compositions such as sub-unit vaccines and inhibitory peptides. Particularly useful are the viral polymerase protein, the spike protein, the nucleocapsid or antigenic fragments thereof for inclusion as antigen or subunit immunogen, but inactivated whole virus can also be used. Particularly useful are also those proteinaceous substances that are encoded by recombinant nucleic acid fragments that are

identified for phylogenetic analyses, of course preferred are those that are within the preferred bounds and metes of ORFs useful in phylogenetic analyses, in particular for eliciting SARS virus specific antibodies, whether in vivo (e.g. for protective puposes or for providing diagnostic antibodies) or in vitro (e.g. by phage display technology or  
5 another technique useful for generating synthetic antibodies).

Also provided herein are antibodies, be it natural polyclonal or monoclonal, or synthetic (e.g. (phage) library-derived binding molecules) antibodies that specifically react with an antigen comprising a proteinaceous molecule or SARS virus-specific functional fragment thereof according to the invention. A person skilled in the art  
10 will be able to develop (monoclonal) antibodies using isolated virus material and/or recombinantly expressed viral proteins. Sui et al. (Proc. Natl. Acad. Sci. 101(8), 2536-2541, 2004) have transiently expressed fragments of the spike protein and found several antibodies through phage display methods. One of these antibodies was shown to be directed to the N-terminal 261-672 amino acids of the S (spike) protein  
15 (which would be corresponding to the sequence of translation 2 of EMC7 and translation 1 of the RDG 1 sequence of the S-protein as depicted in figure 2 ) and this antibody was also demonstrated to have neutralising properties, indicating that it may be a candidate for succesfull vaccines. Also Subbarao *et al.* (supra) showed that serum from mice that had been infected with SARS virus was able to block infectivity  
20 of 100 TCID<sub>50</sub> of SARS virus in Vero cell monolayers, due to the presence of neutralising antibodies.

Such antibodies are also useful in a method for identifying a viral isolate as a SARS virus comprising reacting said viral isolate or a component thereof with an antibody as provided herein. This can for example be achieved by using purified or  
25 non-purified SARS virus or parts thereof (proteins, peptides) using ELISA, RIA, FACS or similar formats of antigen detection assays (Current Protocols in Immunology). Alternatively, infected cells or cell cultures may be used to identify viral antigens using classical immunofluorescence or immunohistochemical techniques. Specifically useful in this respect are antibodies raised against SARS  
30 virus proteins which are encoded by a nucleotide sequence comprising one or more of the fragments disclosed in figure 2.

Other methods for identifying a viral isolate as a SARS virus comprise reacting said viral isolate or a component thereof with a virus specific nucleic acid according to the invention.

In this way the invention provides a viral isolate identifiable with a method according  
5 to the invention as a mammalian virus taxonomically corresponding to a positive-sense single stranded RNA virus identifiable as likely belonging to the SARS virus genus within the family of Coronaviruses.

The method is useful in a method for virologically diagnosing a SARS virus infection of a mammal, said method for example comprising determining in a sample of said  
10 mammal the presence of a viral isolate or component thereof by reacting said sample with a nucleic acid or an antibody according to the invention.

Methods of the invention can in principle be performed by using any nucleic acid amplification method, such as the Polymerase Chain Reaction (PCR; Mullis 1987, U.S. Pat. No. 4,683,195, 4,683,202, en 4,800,159) or by using amplification reactions  
15 such as Ligase Chain Reaction (LCR; Barany 1991, Proc. Natl. Acad. Sci. USA 88:189-193; EP Appl. No., 320,308), Self-Sustained Sequence Replication (3SR; Guatelli et al., 1990, Proc. Natl. Acad. Sci. USA 87:1874-1878), Strand Displacement Amplification (SDA; U.S. Pat. Nos. 5,270,184, en 5,455,166), Transcriptional Amplification System (TAS; Kwok et al., Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-  
20 Beta Replicase (Lizardi et al., 1988, Bio/Technology 6:1197), Rolling Circle Amplification (RCA; U.S. Pat. No. 5,871,921), Nucleic Acid Sequence Based Amplification (NASBA), Cleavase Fragment Length Polymorphism (U.S. Pat. No. 5,719,028), Isothermal and Chimeric Primer-initiated Amplification of Nucleic Acid (ICAN), Ramification-extension Amplification Method (RAM; U.S. Pat. Nos. 5,719,028  
25 and 5,942,391) or other suitable methods for amplification of nucleic acids.

In order to amplify a nucleic acid with a small number of mismatches to one or more of the amplification primers, an amplification reaction may be performed under conditions of reduced stringency (e.g. a PCR amplification using an annealing temperature of 38°C, or the presence of 3.5 mM MgCl<sub>2</sub>). The person skilled in the art  
30 will be able to select conditions of suitable stringency.

The primers herein are selected to be "substantially" complementary (i.e. at least 65%, more preferably at least 80% perfectly complementary) to their target regions present on the different strands of each specific sequence to be amplified. It is possible to use primer sequences containing e.g. inositol residues or ambiguous bases

or even primers that contain one or more mismatches when compared to the target sequence. In general, sequences that exhibit at least 65%, more preferably at least 80% homology with the target DNA or RNA oligonucleotide sequences, are considered suitable for use in a method of the present invention. Sequence mismatches are also  
5 not critical when using low stringency hybridization conditions.

The detection of the amplification products can in principle be accomplished by any suitable method known in the art. The detection fragments may be directly stained or labelled with radioactive labels, antibodies, luminescent dyes, fluorescent dyes, or enzyme reagents. Direct DNA stains include for example intercalating dyes such as  
10 acridine orange, ethidium bromide, ethidium monoazide or Hoechst dyes.

Alternatively, the DNA or RNA fragments may be detected by incorporation of labelled dNTP bases into the synthesized fragments. Detection labels which may be associated with nucleotide bases include e.g. fluorescein, cyanine dye or BrdUrd.

When using a probe-based detection system, a suitable detection procedure for use in  
15 the present invention may for example comprise an enzyme immunoassay (EIA) format (Jacobs et al., 1997, J. Clin. Microbiol. 35, 791-795). For performing a detection by manner of the EIA procedure, either the forward or the reverse primer used in the amplification reaction may comprise a capturing group, such as a biotin group for immobilization of target DNA PCR amplicons on e.g. a streptavidin coated  
20 microtiter plate wells for subsequent EIA detection of target DNA -amplicons (see below). The skilled person will understand that other groups for immobilization of target DNA PCR amplicons in an EIA format may be employed.

Probes useful for the detection of the target DNA as disclosed herein preferably bind only to at least a part of the DNA sequence region as amplified by the DNA  
25 amplification procedure. Those of skill in the art can prepare suitable probes for detection based on the nucleotide sequence of the target DNA without undue experimentation as set out herein. Also the complementary nucleotide sequences, whether DNA or RNA or chemically synthesized analogs, of the target DNA may suitably be used as type-specific detection probes in a method of the invention,  
30 provided that such a complementary strand is amplified in the amplification reaction employed.

Suitable detection procedures for use herein may for example comprise immobilization of the amplicons and probing the DNA sequences thereof by e.g. southern blotting. Other formats may comprise an EIA format as described above. To

facilitate the detection of binding, the specific amplicon detection probes may comprise a label moiety such as a fluorophore, a chromophore, an enzyme or a radio-label, so as to facilitate monitoring of binding of the probes to the reaction product of the amplification reaction. Such labels are well-known to those skilled in the art and include, for example, fluorescein isothiocyanate (FITC),  $\beta$ -galactosidase, horseradish peroxidase, streptavidin, biotin, digoxigenin,  $^{35}\text{S}$  or  $^{125}\text{I}$ . Other examples will be apparent to those skilled in the art.

Detection may also be performed by a so called reverse line blot (RLB) assay, such as for instance described by Van den Brule et al. (2002, J. Clin. Microbiol. 40, 779-787).

For this purpose RLB probes are preferably synthesized with a 5' amino group for subsequent immobilization on e.g. carboxyl-coated nylon membranes. The advantage of an RLB format is the ease of the system and its speed, thus allowing for high throughput sample processing.

The use of nucleic acid probes for the detection of RNA or DNA fragments is well known in the art. Mostly these procedure comprise the hybridization of the target nucleic acid with the probe followed by post-hybridization washings. Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For nucleic acid hybrids, the  $T_m$  can be approximated from the equation of Meinkoth and Wahl, Anal. Biochem., 138: 267-284 (1984):  $T_m = 81.5\text{ }^\circ\text{C} + 16.6 (\log M) + 0.41 (\% \text{ GC}) - 0.61 (\% \text{ form}) - 500/L$ ; where M is the molarity of monovalent cations, % GC is the percentage of guanosine and cytosine nucleotides in the nucleic acid, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe.  $T_m$  is reduced by about 1  $^\circ\text{C}$  for each 1 % of mismatching; thus, the hybridization and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with > 90% identity are sought, the  $T_m$  can be decreased 10 $^\circ\text{C}$ . Generally, stringent conditions are selected to be about 5  $^\circ\text{C}$  lower than the thermal melting point ( $T_m$ ) for the specific sequence and its complement at a defined ionic strength and pH.

However, severely stringent conditions can utilize a hybridization and/or wash at 1,2,3, or 4  $^\circ\text{C}$  lower than the thermal melting point ( $T_m$ ); moderately stringent conditions can utilize a hybridization and/or wash at 6,7,8,9, or 10  $^\circ\text{C}$  lower than the thermal melting point ( $T_m$ ); low stringency conditions can utilize a hybridization

and/or wash at 11,12,13,14,15, or 20 °C lower than the thermal melting point ( $T_m$ ). Using the equation, hybridization and wash compositions, and desired  $T_m$ , those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a  $T_m$  of less than 45 °C (aqueous solution) or 32 °C (formamide solution) it is preferred to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes, Part I, Chapter 2" Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier. New York (1993); and Current Protocols in Molecular Biology, Chapter 2, Ausubel, et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995).

In another aspect, the invention provides oligonucleotide probes for the generic detection of target RNA or DNA. The detection probes herein are selected to be "substantially" complementary to one of the strands of the double stranded nucleic acids generated by an amplification reaction of the invention. Preferably the probes are substantially complementary to the immobilizable, e.g. biotin labelled, antisense strands of the amplicons generated from the target RNA or DNA.

It is allowable for detection probes of the present invention to contain one or more mismatches to their target sequence. In general, sequences that exhibit at least 65%, more preferably at least 80% homology with the target oligonucleotide sequences are considered suitable for use in a method of the present invention. Antibodies, both monoclonal and polyclonal, can also be used for detection purpose in the present invention, for example, in immunoassays in which they can be utilized in liquid phase or bound to a solid phase carrier. In addition, the monoclonal antibodies in these immunoassays can be detectably labeled in various ways. A variety of immunoassay formats may be used to select antibodies specifically reactive with a particular protein (or other analyte). For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See Harlow and Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York (1988), for a description of immunoassay formats and conditions that can be used to determine selective binding. Examples of types of immunoassays that can utilize antibodies of the invention are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such

immunoassays are the radioimmunoassay (RIA) and the sandwich (immunometric) assay. Detection of the antigens using the antibodies of the invention can be done utilizing immunoassays that are run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. Those of skill in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

Antibodies can be bound to many different carriers and used to detect the presence of the target molecules. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding monoclonal antibodies, or will be able to ascertain such using routine experimentation.

The invention also provides a method for serologically diagnosing a SARS virus infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a SARS virus or component thereof by reacting said sample with a proteinaceous molecule or fragment thereof or an antigen according to the invention

Methods and means provided herein are particularly useful in a diagnostic kit for diagnosing a SARS virus infection, be it by virological or serological diagnosis. Such kits or assays may for example comprise a virus, a nucleic acid, a proteinaceous molecule or fragment thereof, an antigen and/or an antibody according to the invention.

Use of a virus, a nucleic acid, a proteinaceous molecule or fragment thereof, an antigen and/or an antibody according to the invention is also provided for the production of a pharmaceutical composition, for example for the treatment or prevention of SARS virus infections and/or for the treatment or prevention of atypical pneumonia, in particular in humans. Preferably a peptide comprising part of the amino acid sequence of the spike protein as depicted in translation 2 with the sequence EMC7 and translation 1 of the RDG seq of figure 2, is used for the preparation of a therapeutic or prophylactic peptide. Also preferably, a protein comprising the amino acid sequence of the spike protein as depicted in translation 2 with the sequence EMC7 translation 1 of the RDG seq of figure 2, is used for the preparation of a sub-unit vaccine. Furthermore, the nucleocapsid of Coronaviruses, as

depicted in the translation of EMC8, in figure 2, is known to be particularly useful for eliciting cell-mediated immunity against Coronaviruses and can be used for the preparation of a sub-unit vaccine.

Attenuation of the virus can be achieved by established methods developed for this purpose, including but not limited to the use of related viruses of other species, serial passages through laboratory animals or/and tissue/cell cultures, serial passages through cell cultures at temperatures below 37°C (cold-adaptation), site directed mutagenesis of molecular clones and exchange of genes or gene fragments between related viruses.

As is shown by Sui et al. (*supra*) humanised neutralising antibodies have been prepared which have shown to be reactive with the N-terminal 261-672 amino acids of the spike protein of the SARS virus.

A pharmaceutical composition comprising a virus, a nucleic acid, a proteinaceous molecule or fragment thereof, an antigen and/or an antibody according to the invention can for example be used in a method for the treatment or prevention of a SARS virus infection and/or a respiratory illness comprising providing an individual with a pharmaceutical composition according to the invention. This is most useful when said individual comprises a human. Antibodies against SARS virus proteins, especially against the spike protein of SARS virus, preferably against the amino acid sequence as depicted in translation 2 of EMC7 and translation 1 of the RDG seq in figure 2, are also useful for prophylactic or therapeutic purposes, as passive vaccines. It is known from other coronaviruses that the spike protein is a very strong antigen and that antibodies against spike protein can be used in prophylactic and therapeutic vaccination.

The invention also provides method to obtain an antiviral agent useful in the treatment of atypical pneumonia comprising establishing a cell culture or experimental animal comprising a virus according to the invention, treating said culture or animal with an candidate antiviral agent, and determining the effect of said agent on said virus or its infection of said culture or animal. An example of such an antiviral agent comprises a SARS virus-neutralising antibody, or functional component thereof, as provided herein, but antiviral agents of other nature are obtained as well. The invention also provides use of an antiviral agent according to the invention for the preparation of a pharmaceutical composition, in particular for the preparation of a pharmaceutical composition for the treatment of atypical



pneumonia, specifically when caused by a SARS virus infection, and provides a pharmaceutical composition comprising an antiviral agent according to the invention, useful in a method for the treatment or prevention of a SARS virus infection or atypical pneumonia, said method comprising providing an individual with such a  
5 pharmaceutical composition.

Specifically the invention provides a pharmaceutical composition comprising interferon, especially pegylated interferon.

In general all interferon forms would be useful in the present invention, since it is known that all the interferon forms have at least some activity in alleviating  
10 (symptoms of) viral infection. However, it is to be understood that preferentially the interferon is used which is derived from the host which is infected with, or which runs the risk of being infected with the virus. Further, most preferred is the use of interferon-alpha, and especially – for coronavirus infections that affect humans, like SARS – human interferon-alpha. Alpha interferon is a natural protein produced by  
15 the human body in response to infection. It is also known as interferon alpha-2b. The type I interferon alpha family consists of small proteins that have clinically important anti-infective and anti-tumor activity. It is understood that alpha interferon may be administered alone or in combination with beta interferon or gamma interferon.

20 Genetic engineering techniques have allowed several companies to mass-produce alpha interferon, which is known as recombinant human alpha interferon, or by abbreviations such as rhIFN or rIFN-alpha. This is marketed under tradenames such as Viraferon (made by Schering-Plough), Roferon-A (by Roche) and Wellferon (by Glaxo SmithKline). Interferon-alpha N3, or Alferon N, is another form of interferon  
25 alpha, derived from human leukocytes and containing multiple species of interferon-alpha.

A drawback to the use of interferon-alpha as discussed previously, is the short serum half-life and rapid clearance of the interferon alpha protein. However, it has been shown that the attachment of molecules of polyethylene glycol (PEG) to the  
30 interferon, creates a barrier that shields the interferon alfa-2a molecule from being rapidly degraded by proteases in the body and maintains its ability to consistently suppress the target virus over a longer dosage period.

As already discussed above pegylation of proteins such as inteferon is used to prevent rapid removal from the bloodstream and eventually rapid breakdown of the

drug. A prolongation of the serum half-life of more than a factor two has been demonstrated (Shannon A. Marshall, Drug Discovery Today Volume 8, Issue 5, March 2003, Pages 212-221). Pegylated IFN alfa-2b has a prolonged serum half-life (40 hours) relative to standard IFN alfa-2b (7-9 hours). The greater size of pegylated IFN alfa-2a acts to reduce glomerular filtration, markedly prolonging its serum half-life (72-96 hours) compared with standard IFN alfa-2a (6-9 hours) (Bruce A. Luxon MD Clinical Therapeutics, Volume 24, Issue 9, September 2002, Pages 1363-1383).

Pegylation of proteins is a standard technique available to a person skilled in the art, and standard pegylated interferons are available commercially from Roche (PEGASYS® (interferon alfa 2a) and Schering-Plough (PEG-Intron A) or in development like PEG-Alfacon, the PEGylated version of Infergen(R) (Interferon alfacon-1) a bio-engineered type I interferon alpha.

Schering-Plough has developed a semi-synthetic form of Intron® A by attaching a 12-kDa mono-methoxy polyethylene glycol to the protein (PEG Intron) which fulfils the requirements of a long-acting interferon alpha protein while providing significant clinical benefits. Pegylation decreases the specific activity of the interferon alpha-2b protein, whilst the potency of PEG Intron, independent of protein concentration, is comparable to the Intron® A standard at both the molecular and cellular level. PEG Intron has enhanced pharmacokinetic profile in both animal and human studies [see Yu-Sen Wang et al., 2002: Advanced Drug Delivery Reviews, Volume 54, Issue 4, 17, Pages 547-570]. In PEGASYS, a 40 kilodalton branched, mobile PEG is covalently bound to the interferon alfa-2a molecule and provides a selectively protective barrier without significantly reducing binding site receptivity.

It is understood that pharmacokinetic behaviour of a pegylated molecule depends on the size of the PEG and the structure of the link between the PEG moiety and the protein (Shannon A. Marshall, Drug Discovery Today Volume 8, Issue 5, March 2003, Pages 212-221). It is known that interferons with smaller PEGs are degraded quickly, requiring more frequent dosing. Thus interferons with larger PEGs are preferred.

Thus the present invention encompasses all types of pegylated interferons or future interferons with yet undisclosed molecule attachments which provide a selectively protective barrier, to shield the interferon from being degraded, without significantly reducing binding site receptivity. Also combinations of different interferons are encompassed in the invention

One of the most preferred embodiments of the present invention is the use of interferon as a prophylactic treatment for the prevention of coronavirus infection. Subjecting apes to a prophylactic or therapeutic treatment either before or during infection with the coronavirus has a good and useful predictionary value for  
5 application of such a prophylaxis or therapy in human subjects.

As is shown in the experimental section administration of interferon before infestation with virus particles greatly delays infection and the effects after infection. It should be understood that the virus challenge given to the test animals is a high  
10 dose, which will not or hardly ever occur in 'natural' infections. It is understood that viral challenge under 'natural' circumstances would equate with a challenge of about 10-105 TCID<sub>50</sub> with a concentration which is much less than that used in the experiment of the invention. Further, the viral challenge in the experiment was for the largest part applied intra-tracheal, i.e. at the place where the virus exerts its  
15 main infective activity. Normally, a virus will be encountered in the air that is breathed and this air will firstly pass the nose and/or oral cavity, where it will have a large chance of being filtered out (and stopped) by the epithelium and mucosa of the mouth and/or the nose. Anyhow, the fact that even at the extremely high dose used in our experiments we have been able to show effect of interferon indicates that the  
20 effect will even be more pronounced at infective viral doses which are normally encountered. It is therefore believed that prophylactic administration will give a durable and strong protection against infection with coronaviruses.

This is especially important in relation to viruses which are highly infective and/or which have an airborne mode of transmission, such as, for instance, the SARS  
25 virus. A prophylactic treatment would be especially welcome for people who run a risk of being infected, such as, in the case of SARS virus, hospital personnel, children, elderly and people having an underlying condition such as diabetes or heart disease, or a weakened immune system.

It remains possible that SARS-CoV infection might be asymptomatic in some  
30 people, or cause nonrespiratory symptoms in others. There is insufficient evidence to exclude the possibility that asymptomatic, or atypical, infected people can transmit the disease. Thus a prophylactic treatment for the prevention of coronavirus infection, like SARS-CoV is indeed essential.

However, our data also show that interferon is also applicable for therapy of coronaviruses, i.e. at the time when virus infection is already established. Our in vivo data show that pathologic effects are at least delayed upon administration of interferon.

5 Interferon of human and murine origins has been quantified in the art in terms of International Units ("IU"). As used herein, a "unit" of interferon (to be distinguished from "IU") shall mean the reciprocal of a dilution of interferon-containing material that, as determined by assay, inhibits one-half the number of plaques of a challenge virus, the challenge virus being the vesicular stomatitis virus  
10 ("VSV"). So quantified a "unit" of interferon is routinely found to be about one-tenth the quantity of interferon represented by one "IU. " Alternatively, interferon can be quantitated in  $\mu\text{g/kg}$  of body weight.

Interferon is given in doses ranging from  $1\mu\text{g/kg}$  to  $3\mu\text{g/kg}$ . When the interferon is pegylated doses can be delivered less frequently. Treatment of a  
15 coronavirus disease in accordance with the present invention comprises administering pegylated interferon at a dosage of  $0.01\text{--}6\mu\text{g/kg}$  per day in a dosage form adapted to promote contact of said dosage of interferon with the oral and pharyngeal mucosa of said animal. Preferably, the dosage of interferon is from  $0.1\text{--}4\mu\text{g/kg}$  per day, more preferably  $0.3\text{--}3\mu\text{g/kg}$  per day.

20 Interferon may be administered by any available means, including but not limited to, oral, intravenous, intramuscular, pulmonary and nasal routes, and wherein said composition is present as a solution, a suspension or an aerosol spray, especially of fine particles.

It is critical that the pegylated interferon be administered in a dosage form  
25 adapted to assure maximum contact of the interferon in said dosage form with the oral and pharyngeal mucosa of the human or animal, undergoing treatment. Contact of interferon with the mucosa can be enhanced by maximizing residence time of the treatment solution in the oral or pharyngeal cavity. Thus, best results seem to be achieved in human patients when the patient is requested to hold said solution of  
30 interferon in the mouth for a period of time. Contact of interferon with the oral and pharyngeal mucosa and thereafter with the lymphatic system of the treated human or animal avian, rodent is unquestionably the most efficient method administering immunotherapeutic amounts of pegylated interferon.

For example interferon can be administered in either a liquid (solution) or solid dosage form. Thus interferon can be administered dissolved in a buffered aqueous solution typically containing a stabilizing amount (1-5% by weight) of blood serums. Exemplary of a buffered solution suitable as a carrier of interferon  
5 administered in accordance with this invention is phosphate buffered saline prepared by standard techniques.

It is also contemplated by the present invention to provide interferon in a solid dosage form such as a lozenge adapted to be dissolved upon contact with saliva in the mouth with or without the assistance of chewing. Such a unitary dosage form is  
10 formulated to release about 1 to about 1500 IU of interferon upon dissolution in the mouth for contact with the oral and pharyngeal mucosa. Thus a unitary dosage form of interferon in accordance with this invention can be prepared by art-recognized techniques for forming compressed tablets such as chewable vitamins. Similarly, interferon can be incorporated into starch-based gel formulations to form a lozenge  
15 which will dissolve and release interferon for contact with the oral mucosa when held in the mouth. Solid unitary dosage forms of interferon for use in accordance with the present invention can be prepared utilizing art recognized dosage formulation techniques. The pH of such formulations can range from about 4 to about 8.5. Of course, in processing to such unitary dosage forms one should avoid heating a pre-  
20 dosage form formulation, after addition of interferon, above about 50°C. Exemplary of a solid dosage form for animal use is a molasses block containing effective amounts of interferon.

Alternatively the interferon can be formulated into flavoured or unflavoured solutions or syrups using a buffered aqueous solution of interferon as a base with added caloric  
25 or non-caloric sweeteners, flavour oils and pharmaceutically acceptable surfactant/dispersants.

Also contemplated are methods of gene therapy capable of causing expression of interferon in respiratory or gastric cells for prevention of a coronaviral infection.

Of course, the clinical use of any medicament of the present invention is a  
30 clinical decision to be made by the clinician and the exact course of such treatment is left to the clinician's sound discretion, with all such courses of treatment deemed within the bounds of the present invention.

Another preferred embodiment is administration of interferon together with another treatment which is directed to prevent or treat infection with coronaviruses.

Such other treatment can for instance be a vaccine, antibody and/or anti-viral agent selected from the group consisting of whole inactivated virus vaccines, attenuated vaccines, sub-unit vaccines, recombinant vaccines, antibody for passive immunization, nucleoside analogs such as ribavirin, RNA-dependent RNA

5 polymerase inhibitors and protease inhibitors.

Use of interferon together with administration of a vaccine will boost the effects of the vaccine. First of all, there is the combination of treatments that will add up to a better effect. However, co-administration of interferon with a vaccine will also enable the immune response to vaccination to have more effect. Normally the  
10 immune response is slow and it takes a few days to come to a high enough titer of antibodies to be able to effectively combat virus particles. When no interferon is co-administered the virus would have had the chance to multiply to enormous amounts, which cannot be overcome by the immune response. With interferon, however, the amounts of the virus will remain absent or low and any infective virus outburst (if  
15 any at all) can easily be handled by the immune system.

Treatment to prevent and/or treat infection with coronaviruses can also comprise combination treatments with other antiviral compounds, such as, for instance, nucleoside-based compounds such as ribavirin (e.g. Rebetol® (ribavirin, USP). These compounds act through interfering with the viral replication by  
20 presenting nucleosides which are built in during viral replication, but which either prevent formation of viral proteins or which do not yield functional proteins. Co-administration of interferon will even more slow down viral replication. Combinations of ribavirin and forms of interferon can help to reduce viral load.

Another disease condition responding to treatment in accordance with the  
25 present invention is neoplastic disease. Thus, the administration of interferon in accordance with the above description can, alone or in combination with other drugs or therapy, help effect remission of cancers such as malignant lymphoma, melanoma, mesothelioma, Burkitt lymphoma and nasopharyngeal carcinoma and other neoplastic diseases, especially those of known or suspected viral etiology and diseases  
30 such as Hodgkin's Disease and leukemia.

Other disease conditions responding to treatment in accordance with the present invention are infectious diseases of coronaviral origin in human, avian, porcine, canine and feline species. Human coronaviruses are coronavirus 229E and the newly discovered coronavirus HCoV-NL (see European patent application 03078772.5.

Several other coronaviruses can cause fatal systemic diseases in animals, including feline infectious peritonitis virus (FIPV), hemagglutinating encephalomyelitis virus (HEV) of swine, and some strains of avian infectious bronchitis virus (IBV) and mouse hepatitis virus (MHV). These coronaviruses can replicate in liver, lung,  
5 kidney, gut, spleen, brain, spinal cord, retina, and other tissues. Immunopathology plays a role in tissue damage in MHV and FIPV, and cytokines are responsible for some signs of disease. Significantly, in cats with persistent, inapparent infection with feline enterotropic coronavirus, virulent virus mutants can arise and cause fatal infectious peritonitis, a systemic disease.

10

The invention also comprises an animal model usable for testing of prophylactic and/or therapeutic methods and/or preparations. It has appeared that apes can be infected with the SARS virus, thereby showing clinical symptoms, and more importantly, similar tissue morphology as found in humans suffering from  
15 atypical pneumonia caused by the SARS virus. Subjecting apes to a prophylactic or therapeutic treatment either before or during infection with the virus will have a good and useful predictionary value for application of such a prophylaxis or therapy in human subjects.

The invention is further explained in the Examples without limiting it thereto.

## Figure legends

Fig. 1: Phylogenetic relationship for the nucleotide sequences of isolate HK39849 with its closest relatives genetically. Phylogenetic trees were generated by maximum likelihood analyses using 100 bootstraps and 3 jumbles. The scale representing the number of nucleotide changes is shown for each tree.

Fig. 2: Nucleotide sequences from 13 clones of parts of the SARS virus. Also included are the putative polypeptide sequences of polypeptides and alignments of the putative polypeptides with that of another member of the Coronaviridae family, where possible.

Fig. 3: Schematic map of the SARS virus genome, indicating the position of the nucleotide sequences of figure 2 relative to the genome and a putative indication of the open reading frames of the genome based on analogy with other coronaviruses. The gene structure for the region between the Spike and Nucleocapsid is uncertain. EMC1-EMC14 and RDG 1: sequences as provided in figure 2. CDC and BIN1-2: sequences were provided through personal communication from the CDC (Dr. W. Bellini, Centers for Disease Control & Prevention, National Centers for Infectious Diseases, 1600 Clifton Road, Atlanta GA 30333, USA) and BNI (Dr. C. Drosten and Prof. Dr. H. Schmitz, Bernard Nocht Institute, Bernard-Nocht Str. 74, D-20359 Hamburg, Germany), respectively.

Fig. 4: Amino acid comparison of the N-terminus of the S-protein of the SARS virus and closely related coronaviruses. HCV OC43 = human coronavirus isolate OC43; MHV A59 = murine hepatitis virus isolate A59, BCV = bovine corona virus.

Fig. 5: Negative contrast EM photograph of SARS virus obtained from concentrated supernatant of infected cell cultures.

30

Fig. 6: Infection with SARS-coronavirus causes pulmonary and renal lesions in cynomolgus macaques. Formalin-fixed, paraffin-embedded tissues were stained with haematoxylin and eosin and examined by light microscopy. There is diffuse alveolar damage of the lung (a), and the alveolar lumina (b) are flooded with highly



proteinaceous exudate admixed with inflammatory cells and cellular debris. In the lumen of a bronchiole (c) and in the surrounding lung parenchyma are several multinucleated syncytial cells (arrowheads). The renal collecting tubules (d) contain similar multinucleated syncytial cells. Original magnifications: a x 12.5; b x 50; c x 100; d x 250.

Fig 7: Infection of domestic cats and ferrets with SCV. Cats (A, n=6) and ferrets (B, n=6) were infected with  $10^6$  TCID<sub>50</sub> via the respiratory route and secretion of SCV in pharyngeal swabs was quantified by real time PCR. Four animals per group were euthanised at day 4 while the other two were analysed till day 28. SCV secretion in non-infected cats (C, n=2) and non-infected ferrets (D, n=2) exposed to SCV infected cats and ferrets, respectively. Real time PCR results are shown relative to a titrated SCV standard and shown as TCID<sub>50</sub>/ml (N.D. not done).

Fig. 8: Detection of SCV in postmortem tissues of experimentally SCV infected cats and ferrets

Fig. 9: Effect of pegylated IFN- $\alpha$  on SARS Coronavirus (SCV) replication in macaques. SCV detection in pharyngeal swabs (days 0, 2 and 4 after infection, closed bars) and lungs (day 4, open bars) taken from cynomolgus monkeys treated with PBS (A), PEG-Intron at days -3, -1, +1 and +3 (B and C) and PEG-Intron at days +1 and +3 (D) after SCV infection. Individual macaques are shown (n=2 per group). Virus isolation (VI) results are indicated in the lower part of the panel whereas real time PCR results are shown in the upper part of the panels (n.a., not available).

Fig. 10 Nucleotide sequence of SARS Corona virus Genbank accession nr. AY274119

Fig. 11 Antiviral activity of pegylated IFN- $\alpha$  against SCV in vitro and its biological activity in cynomolgus macaques. (a) Effect of pegylated IFN- $\alpha$  against SCV infection in vitro. Similar results were obtained in 3 separate experiments. (b) Pharmacokinetic analysis of pegylated IFN- $\alpha$  in macaques treated with PBS (control group; open squares, n = 4) or pegylated IFN- $\alpha$  (prophylactic group; closed squares, n = 4) at days -3 and -1. (c) Induction of neopterin in macaques treated with PBS

(control group; open squares,  $n = 4$ ) or pegylated IFN- $\alpha$  (prophylactic group; closed squares,  $n = 4$ ) at days  $-3$  and  $-1$ . Data are expressed as mean  $\pm$  s.d.; \*\*,  $P < 0.01$  versus control.

- 5 Fig. 12 Effect of pegylated IFN- $\alpha$  on SCV excretion in cynomolgus macaques. SCV detection in pharyngeal swabs taken at 0, 2, or 4 d.p.i. from macaques treated with PBS (control group,  $n = 4$ ), pegylated IFN- $\alpha$  prophylactically ( $n = 6$ ) or post-exposure ( $n = 4$ ). Data are expressed as mean  $\pm$  s.d.; \*,  $P < 0.05$  versus control group at 2 d.p.i., \*\*,  $P < 0.01$  versus control group at 2 d.p.i..

10

- Fig. 13 Effect of pegylated IFN- $\alpha$  on SCV replication, viral antigen expression and histological lesions in the lungs of SCV-infected cynomolgus macaques. (a) SCV titration of lung homogenates. (b) Immunohistochemical detection of SCV-infected cells in lung sections. (c) Histopathological score of lung sections. SCV-infected  
15 macaques were treated with pegylated IFN- $\alpha$  prophylactically ( $n = 4$ ) or post-exposure ( $n = 4$ ), or treated with PBS (control group,  $n = 4$ ). Data are expressed as mean  $\pm$  s.d.; \*,  $P < 0.05$  versus control group; \*\*,  $P < 0.01$  versus control group.

## Examples

### Example 1. Virus isolation and characterisation

#### *Isolation*

5 Isolate HK39849 was isolated from a hospitalised SARS patient by throat swab and inoculated into a culture of Vero-E6 cells. A sample of the supernatant from these infected cells provided by Dr. M. Peiris (Queen Mary Hospital Faculty of Medicine, Hong Kong University, Honk Kong) was used to inoculate VERO-118 cells and cell culture supernatant from these cells was aliquoted and frozen after one  
10 passage.

We isolated RNA from the virus-containing cell culture supernatant and subjected it to RNA arbitrarily primed PCR (RAP-PCR) essentially as described by Welsh & McClelland (NAR 18:7213; PNAS USA 90:10710, 1993). Virus in the culture supernatants was purified on continuous 20-60% sucrose gradients. The gradient  
15 fractions were inspected for virus-like particles by EM, and RNA was isolated from the fraction containing , in which the most nucleocapsids were observed. Equivalent amounts of RNA isolated from virus fractions were used for RAP-PCR, after which samples were run side by side on a 3% NuSieve agarose gel. Differentially displayed bands ranging in size from 200-1500 base pairs specific for the unidentified virus  
20 were subsequently purified from the gel, cloned in plasmid pCR2.1 (Invitrogen) and sequenced with vector-specific primers. When we used these sequences to search for homologies against sequences in the Genbank database using the BLAST software ([www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)) which yielded resemblance to virus sequences of the coronaviruses displayed in the phylogenetic tree of figure 1.

25 Eight of these fragments (EMC 1-6, 13 and 14) were located in the ORF coding for the viral polymerase (ORF 1ab), one (EMC-7) spanned the 3' end of ORF1ab and reached into the 5' end of spike protein region; EMC-10 overlapped the 3' end of EMC-7 and therefore also codes part of the S protein region and EMC 9 encodes a region downstream of EMC-10; by use of primers to sequences within EMC10 and EMC9  
30 (see below), the region between these two sequences was amplified by PCR and sequenced. The full contiguous region has been incorporated into EMC7 in figure 2; a further sequence (RDG1 in figure 2) encodes the 3' end of the Spike protein. A further sequence (EMC8) spanned part of the Nucleocapsid coding sequence. The remaining three sequences (EMC9, 11 and 12) have in the meantime been found to be regions of

the orf 1ab/replicase, where emc 9 is incorporated in emc 11. This has not yet been reflected in figure 3.

### *Phylogeny*

5 BLAST searches using nucleotide sequences obtained from the unidentified virus isolate revealed homologies primarily with members of the Coronaviridae. As an indication for the relation between the newly identified virus isolate and other coronaviruses a phylogenetic tree was constructed based on the sequence information obtained (figure 1).

10

### *Materials and Methods*

#### *Specimen collection*

Virus was collected from SARS patients using throat swabs and from  
15 experimentally infected monkeys (throat and nasal swabs, serum, plasma and faeces)

#### *Virus isolation and culture*

Throat swabs were dipped into a culture of Vero-E6 cells and incubated for 1-4 days. Cell culture supernatant was clarified by centrifugation and filtered through a  
20 0.45micrometre filter, before beings stored frozen. The virus was subsequently propagated in Vero-118 cells.

#### *Antigen detection by indirect IFA*

Samples from experimentally infected monkeys was cultured on Vero-118 cells  
25 in 24 well plates containing glass slides. These glass slides were washed with PBS and fixed in acetone for 1 minute at room temperature. After washing with PBS the slides were incubated for 30 minutes at 37 °C with SARS-antibody containing serum from SARS patients. After washing off the human serum in PBS, the slides were incubated at 37°C for 30 minutes with FITC labeled anti-human antibodies. After  
30 three washes in PBS and one in tap water, the slides were included in a glycerol/PBS solution (Citifluor, UKC, Canterbury, UK) and covered. The slides were analysed using an Axioscop fluorescence microscope (Carl Zeiss B.V., Weesp, the Netherlands).

35 *Detection of antibodies in humans by indirect IFA*

Virus was cultured on Vero-118 cells in 24 well plates containing glass slides. These glass slides were washed with PBS and fixed in acetone for 1 minute at room temperature. After washing with PBS the slides were incubated for 30 minutes at 37 °C with SARS-antibody containing serum from SARS patients. After washing off the human serum in PBS, the slides were incubated at 37°C for 30 minutes with FITC labeled anti-human antibodies. After three washes in PBS and one in tap water, the slides were included in a glycerol/PBS solution (Citifluor, UKC, Canterbury, UK) and covered. The slides were analysed using an Axioscop fluorescence microscope (Carl Zeiss B.V., Weesp, the Netherlands)

#### *Detection of antibodies in humans by ELISA*

##### *Patient samples.*

4 samples of patients with SARS disease , 8 samples of patients from routine serological virology; samples from an experimentally infected monkey (preserum, 9 and 12 days after infection ).

##### *The Conjugate.*

Whole virus was used as the conjugate.. Tissue culture supernatant from infected Vero cells were pelleted through 20% sucrose onto a 60% sucrose cushion. The virus was then pelleted through 20% sucrose and resuspended in PBS/1% NP40. After dialysis using PBS, the virus was The conjugated to horseradish peroxidase by standard techniques was tested in 3 concentrations (diluted in dilution buffer 9000-03, 1:100, 1:400 and 1:1600), both on polyvalent anti-IgM code MCB0201 (cross-reactive with monkey) and monoclonal anti-IgM, code 9000-62 (non-crossreactive with monkey).

Sera were diluted 1:200 in serum diluent (code 9000-03), monkey 775 was diluted 1: 100, 1:200 and 1:400.

Serum incubation one hour at 37°C, conjugate incubation one hour at 37°C, and TMB (ready to use): 30 minutes at room temperature. The reaction was stopped with sulphuric acid (0.5M).

*Virus characterisation*

For EM analyses, virus was concentrated from infected cell culture supernatants in a micro-centrifuge at 4 °C at 17000 x g, after which the pellet was  
5 resuspended in PBS and inspected by negative contrast EM

*RNA isolation*

RNA was isolated from the supernatant of infected cell cultures or sucrose  
10 gradient fractions using a High Pure RNA Isolation kit according to instructions from the manufacturer (Roche Diagnostics, Almere, The Netherlands).

*RT-PCR*

A one-step RT-PCR was performed in 50 µl reactions containing 50 mM  
15 Tris.HCl pH 8.5, 50 mM NaCl, 4 mM MgCl<sub>2</sub>, 2 mM dithiothreitol, 200 µM each dNTP, 10 units recombinant RNasin (Promega, Leiden, the Netherlands), 10 units AMV RT (Promega, Leiden, The Netherlands), 5 units Amplitaq Gold DNA polymerase (PE Biosystems, Nieuwerkerk aan de IJssel, The Netherlands) and 5 µl RNA. Cycling conditions were 45 min. at 42 °C and 7 min. at 95 °C once, 1 min at 95 °C, 2 min. at  
20 42 °C and 3 min. at 72 °C repeated 40 times and 10 min. at 72 °C once.

Primers used for diagnostic PCR:

SARS fwd2: ggtggaacatcatccggtgat

SARS rev2: agcctgtgttagattgcgg

These primers amplify a 149bp fragment of the polymerase gene (orf 1ab)

25 RF 999: TTAAACACTTACGAGAGTTTGTG

RF997: GGACACAACCCATGAAATCATCTGG

These primers amplify a region of 728bp in the spike glycoprotein gene (S)

RF998: AGACATATCTAATGTGCCTTTCTCC RF1002:

AAGCTCGTCACCTAAGTCATAAGAC (from EMC11 sequence)

30 The combination of RF998/RF1002 primers enabled us to sequence the 3' end of EMC7 – RF998 is a specific primer withing EMC7 whereas EMC1002 acted as a random primer.

RT-PCR, gel purification and direct sequencing were performed as described above.

### *RAP-PCR*

RAP-PCR was performed essentially as described by Welsh & McClelland (Nuc. Acid  
5 Res. 18:7213, 1990; Proc. Natl. Acad. Sci. USA 90:10710 1993) . The oligonucleotide  
sequences are described in addenda 2. For the RT reaction, 2 µl RNA was used in a  
10 µl reaction containing 10 ng/µl oligonucleotide, 10 mM dithiotreitol, 500 µM each  
dNTP, 25 mM Tris-HCl pH 8.3, 75 mM KCl and 3 mM MgCl<sub>2</sub>. The reaction mixture  
was incubated for 5 min. at 70 °C and 5 min. at 37 °C, after which 200 units  
10 Superscript RT enzyme (LifeTechnologies) were added. The incubation at 37 °C was  
continued for 55 min. and the reaction terminated by a 5 min. incubation at 72 °C.  
The RT mixture was diluted to give a 50 µl PCR reaction containing 8 ng/µl  
oligonucleotide, 300 µM each dNTP, 15 mM Tris-HCL pH 8.3, 65 mM KCl, 3.0 mM  
MgCL<sub>2</sub> and 5 units Taq DNA polymerase (PE Biosystems). Cycling conditions were 5  
15 min. at 94 °C, 5 min. at 40 °C and 1 min. at 72 °C once, followed by 1 min. at 94 °C, 2  
min. at 56 °C and 1 min. at 72 °C repeated 40 times and 5 min. at 72°C once. After  
RAP-PCR, 15 µl the RT-PCR products were run side by side on a 3% NuSieve agarose  
gel (FMC BioProducts, Heerhugowaard, The Netherlands). Differentially displayed  
fragments were purified from the gel with Qiaquick Gel Extraction kit (Qiagen,  
20 Leusden, The Netherlands) and cloned in pCR2.1 vector (Invitrogen, Groningen, The  
Netherlands) according to instructions from the manufacturer.

### *Sequence analysis*

RAP-PCR products cloned in vector pCR2.1 (Invitrogen) were sequenced with M13-  
25 specific oligonucleotides. DNA fragments obtained by RT-PCR were purified from  
agarose gels using Qiaquick Gel Extraction kit (Qiagen, Leusden, The Netherlands),  
and sequenced directly with the same oligonucleotides used for PCR. Sequence  
analyses were performed using a Dyenamic ET terminator sequencing kit  
(Amersham Pharmacia Biotech, Roosendaal, The Netherlands) and an ABI 373  
30 automatic DNA sequencer (PE Biosystem). All techniques were performed according  
to the instructions of the manufacturer.

### *RT-PCR for diagnosing SARS virus.*

For the amplification of the SARS virus' genetic material, we used primers:

SARS fwd2: ggtggaacatcatccggtgat

SARS rev2: agcctgtgttagattgcgg

These primers amplify a 149bp fragment of the polymerase gene (orf 1ab)

RF 999: TTAAACACTTACGAGAGTTTGTG

5 RF997: GGACACAACCCATGAAATCATCTGG

These primers amplify a region of 728bp in the spike glycoprotein gene (S)

These primers amplify a 149bp fragment of the polymerase gene (orf 1ab)

RT-PCR, gel purification and direct sequencing were performed as described above.

10

### *Phylogenetic analyses*

For all phylogenetic trees, DNA sequences were aligned using the ClustalW software package and maximum likelihood trees were generated using the DNA-ML software  
15 package of the Phylip 3.5 program using 100 bootstraps and 3 jumbles<sup>15</sup>. Previously published sequences for TGEV, PEDV, 229E, AIBV, BoCo and MHV that were used for the generation of phylogenetic trees are available from Genbank

### **Example 2: Methods to identify SARS virus**

#### 20 *Specimen collection*

In order to find virus isolates nasopharyngeal aspirates, throat and nasal swabs, broncho alveolar lavages, serum and plasma samples, and stools preferably from mammals such as humans, carnivores (dogs, cats, mustellids, seals etc.), horses,  
25 ruminants (cattle, sheep, goats etc.), pigs, rabbits, birds (poultry, ostriches, etc) should be examined. From birds cloaca swabs and droppings can be examined as well. Sera should be collected for immunological assays, such as ELISA, molecular-based assays, such as RT-PCR and virus neutralisation assays.

Collected virus specimens were diluted with 5 ml Dulbecco MEM medium  
30 (BioWhittaker, Walkersville, MD) and thoroughly mixed on a vortex mixer for one minute. The suspension was thus centrifuged for ten minutes at 840 x g. The sediment was spread on a multispot slide (Nutacon, Leimuiden, The Netherlands) for immunofluorescence techniques, and the supernatant was used for virus isolation.



*Virus isolation*

For virus isolation Vero-118 cells or tMK cells (RIVM, Bilthoven, The Netherlands) were cultured in 24 well plates containing glass slides (Costar, Cambridge, UK), with  
5 the medium described below supplemented with 10% fetal bovine serum (BioWhittaker, Vervier, Belgium). Before inoculation the plates were washed with PBS and supplied with Eagle's MEM with Hanks' salt (ICN, Costa mesa, CA) supplemented with 0.52/liter gram  $\text{NaHCO}_3$ , 0.025 M Hepes (Biowhittaker), 2 mM L-glutamine (Biowhittaker), 200 units/liter penicilline, 200  $\mu\text{g/liter}$  streptomycine  
10 (Biowhittaker), 1gram/liter lactalbumine (Sigma-Aldrich, Zwijndrecht, The Netherlands), 2.0 gram/liter D-glucose (Merck, Amsterdam, The Netherlands), 10 gram/liter peptone (Oxoid, Haarlem, The Netherlands) and 0.02% trypsin (Life Technologies, Bethesda, MD). The plates were inoculated with supernatant of the patient samples, 0,2 ml per well in triplicate, followed by centrifuging at 840x g for  
15 one hour. After inoculation the plates were incubated at 37 °C for a maximum of 1-3 days and cultures were checked daily for CPE. Extensive CPE was generally observed within 24hours. and included detachment of cells from the monolayer..

*Virus culture of SARS*

20 Sub-confluent monolayers of tMK cells or Vero clone 118 cells in media as described above were inoculated with supernatants of samples that displayed CPE or with samples taken from patient or artificially infected monkeys..

*Virus characterisation*

25 For EM analyses, virus was concentrated from infected cell culture supernatants in a micro-centrifuge at 4 °C at 17000 x g, after which the pellet was resuspended in PBS and inspected by negative contrast EM.

*Antigen detection by indirect IFA*

30 Virus was cultured on Vero-118 cells in 24 well slides containing glass slides. These glass slides were washed with PBS and fixed in acetone for 1 minute at room temperature.  
After washing with PBS the slides were incubated for 30 minutes at 37 °C with SARS patient serum. We used patient serum, but antibodies can be raised in various

animals, such as ferrets, goats and rabbits (for polyclonal antibodies) and mice and hamsters (for monoclonal antibodies), and the working dilution of the antibody can vary for each immunisation. After three washes with PBS and one wash with tap water, the slides were incubated at 37°C for 30 minutes with FITC labeled goat-anti-  
5 human antibodies. After three washes in PBS and one in tap water, the slides were included in a glycerol/PBS solution (Citifluor, UKC, Canterbury, UK) and covered. The slides were analysed using an Axioscop fluorescence microscope (Carl Zeiss B.V., Weesp, the Netherlands).

10 *Detection of antibodies in humans by indirect IFA*

For the detection of virus specific antibodies, SARS virus-infected Vero cells were fixed with acetone on coverslips (as described above), washed with PBS and incubated 30 minutes at 37°C with serum samples at a 1 to 16 dilution. After two washes with PBS and one with tap water, the slides were incubated 30 minutes at  
15 37°C with FITC-labelled secondary antibodies to human antibodies (Dako). Slides were processed as described above.

Antibodies can be labelled directly with a fluorescent dye, which will result in a direct immuno fluorescence assay. FITC can be replaced with any fluorescent dye. This technique can be applied to antibodies in other animals such as mammals,  
20 ruminants, birds or other species, assuming the secondary antibody to the appropriate species is used.

*Detection of antibodies in humans by ELISA*

*Patient samples.*

25 4 samples of patients with SARS; 8 samples of patients from routine serological virology; samples from an experimentally infected monkey (preserum and 9 days after infection).

*The Conjugate.*

30 The conjugate was tested at a number of concentrations, both on polyvalent anti-IgM (cross-reactive with monkey) and monoclonal anti-IgM, (non-crossreactive with monkey).

Sera were diluted 1:200 in serum diluent and the monkey serum was diluted 1: 100, 1:200 and 1:400.

Serum incubation one hour at 37°C, conjugate incubation one hour at 37°C, and TMB (ready to use): 30 minutes at room temperature. The reaction was stopped with  
5 sulphuric acid (0.5M).

Results were interpreted by eye. Three of the four SARS-IgM positive sera (as detected by IF on infected cells) had a higher score than negative control sera. One serum had a score which was also reached by some of the negative controls. The 9 day  
10 old monkey sera did not react, but the 12 day old did. Thus, this study shows that with direct conjugation of nucleocapsids the developemnt of an IgM capture method is feasible.

Furthermore, this type of assay can be performed in a number of formats by those trained in the art. The assay can be extended to the detection of IgA and IgG  
15 antibodies from humans and animals and can make use of different capture antigens, such as, but not limited to, purified recombinant N protein.

#### *Animal immunisation*

Cynomologous macaque specific antisera for the newly discovered virus were  
20 generated by experimental intratracheal installation of cultured virus of Cynomologous macaques. One to two weeks later the animals were bled. The sera were tested for reactivity to SARS virus by indirect IFA as described above; uninfected control cells were used to ensure the specificity of the serum. Other animal species are also suitable for the generation of specific antibody preparations  
25 and other antigen preparations may be used.

#### *RNA isolation*

RNA was isolated from the supernatant of infected cell cultures or sucrose  
30 gradient fractions using a High Pure RNA Isolation kit according to instructions from the manufacturer (Roche Diagnostics, Almere, The Netherlands). RNA can also be isolated following other procedures known in the field (*Current Protocols in Molecular Biology*).

*RT-PCR*

A one-step RT-PCR was performed in 50 µl reactions containing 50 mM Tris.HCl pH 8.5, 50 mM NaCl, 4 mM MgCl<sub>2</sub>, 2 mM dithiotreitol, 200 µM each dNTP, 10 units recombinant RNAsin (Promega, Leiden, the Netherlands), 10 units AMV RT (Promega, Leiden, The Netherlands), 5 units Amplitaq Gold DNA polymerase (PE Biosystems, Nieuwerkerk aan de IJssel, The Netherlands) and 5 µl RNA. Cycling conditions were 45 min. at 42 °C and 7 min. at 95 °C once, 1 min at 95 °C, 2 min. at 42 °C and 3 min. at 72 °C repeated 40 times and 10 min. at 72 °C once.

Primers used for diagnostic PCR:

10

For the amplification of the SARS virus' genetic material, we used primers:

SARS fwd2: ggtggaacatcatccggtgat

SARS rev2: agcctgtgtgtagattgegg

These primers amplify a 149bp fragment of the polymerase gene (orf 1ab)

15 RT-PCR, gel purification and direct sequencing were performed as described above.

*Sequence analysis*

Sequence analyses were performed using a Dyenamic ET terminator sequencing kit (Amersham Pharmacia Biotech, Roosendaal, The Netherlands) and an ABI 373 automatic DNA sequencer (PE Biosystem). All techniques were performed according to the instructions of the manufacturer. PCR fragments were sequenced directly with the same oligonucleotides used for PCR, or the fragments were purified from the gel with Qiaquick Gel Extraction kit (Qiagen, Leusden, The Netherlands) and cloned in pCR2.1 vector (Invitrogen, Groningen, The Netherlands) according to instructions from the manufacturer and subsequently sequenced with M13-specific oligonucleotides.

25

*Detection of antibodies in humans, mammals, ruminants or other animals by ELISA*

A recombinant protein derived from the SARS virus is preferred as the antigen. However, purified nucleocapsids may also be used. Antigens suitable for antibody detection include any SARS protein that combines with any SARS-specific antibody of a patient exposed to or infected with SARS virus. Preferred antigens of the invention include those that predominantly engender the immune response in

30

patients exposed to SARS, which therefore, typically are recognised most readily by antibodies of a patient. Particularly preferred antigens include the N, and S proteins of SARS.

Antigens used for immunological techniques can be native antigens or can be modified versions thereof. Well known techniques of molecular biology can be used to alter the amino acid sequence of a SARS antigen to produce modified versions of the antigen that may be used in immunologic techniques.

Methods for cloning genes, for manipulating the genes to and from expression vectors, and for expressing the protein encoded by the gene in a heterologous host are well-known, and these techniques can be used to provide the expression vectors, host cells, and the for expressing cloned genes encoding antigens in a host to produce recombinant antigens for use in diagnostic assays. See for instance: *Molecular cloning, A laboratory manual* and *Current Protocols in Molecular Biology*.

A variety of expression systems may be used to produce SARS antigens. For instance, a variety of expression vectors suitable to produce proteins in *E.Coli*, *B.subtilis*, yeast, insect cells and mammalian cells have been described, any of which might be used to produce a SARS antigen suitable to detect anti- SARS antibodies in exposed patients.

The baculovirus expression system has the advantage of providing necessary processing of proteins, and is therefor preferred. The system utilizes the polyhedrin promoter to direct expression of SARS antigens. (Matsuura et al. 1987, J.Gen.Virol. 68: 1233-1250).

Antigens produced by recombinant baculo-viruses can be used in a variety of immunological assays to detect anti- SARS antibodies in a patient. It is well established, that recombinant antigens can be used in place of natural virus in practically any immunological assay for detection of virus specific antibodies. The assays include direct and indirect assays, sandwich assays, solid phase assays such as those using plates or beads among others, and liquid phase assays. Assays suitable include those that use primary and secondary antibodies, and those that use antibody binding reagents such as protein A. Moreover, a variety of detection methods can be used in the invention, including colorimetric, fluorescent, phosphorescent, chemiluminescent, luminescent and radioactive methods.

**Example 3: Animal models***Macaques*

Four Cynomologous Macaques were infected with SARS virus by intratracheal installation using Vero-118 cell derived virus.

5

The monkeys had the following clinical symptoms

- Lethargy
- One of four monkeys had severe pneumonia
- Mild to severe rash in the inguinal region and the axilar region
- 10 ◦ Watery stools

After 10-16 days the monkeys were euthanized. Tissues were examined and the following was found

- Alveolae were filled with serum and their architecture were disrupted,  
15 consistent with bronchointestinal pneumonia (see fig 5 and b)
- Multi-cell syncytia in lungs (fig 5c)
- Multi-cell syncytia in kidneys (fig 5d)
- Widening of the small intestine

20 Virus was detected using RT-PCR on tissue samples and by culturing samples followed by electron microscopy from

- Lungs
- Nasal swabs
- Throat swabs
- 25 • Faeces
- Kidneys

The EM results demonstrate that the virus that was recovered from the Cynomologous Macaques had the identical morphology to the virus which was used to infect them.

30

This demonstrates that Cynomologous Macaques may be used as animal models to tests the efficacy of pharmaceutical preparations for therapeutic or prophylactic purposes

*Cats and ferrets*

Domestic cats (n = 6) and ferrets (n = 6) were inoculated intratracheally with 106 median tissue culture infectious dose (TCID<sub>50</sub>) SCV, obtained from patient 5688 who died from SARS and passaged four times on Vero 118 cells in vitro. Nasal, pharyngeal and rectal swabs were taken on different days post infection (d.p.i.). Four animals of each group were euthanised at 4 d.p.i. and necropsy was performed according to a standard protocol. No clinical signs were observed in SCV-inoculated cats, while three out of six ferrets became lethargic from 2 to 4 d.p.i. and one of these ferrets died at 4 d.p.i. All cats and ferrets (Fig. 7) shed SCV from the pharynx starting at 2 d.p.i. until day 10 and 14, respectively, as determined by RT-PCR. Virus was isolated from all pharyngeal swabs taken on 2-8 d.p.i. and nasal swabs of two cats on 4 and 6 d.p.i.. SCV was detected neither in nasal swabs from ferrets nor in rectal swabs of cats or ferrets. Infection of the respiratory tract was evident in all animals tested; SCV could be isolated from their tracheas and lungs (Fig 8). Quantification of the mean geometric viral titres per ml lung homogenate revealed relatively low SCV titers in the lungs of SCV-inoculated cats ( $1 \times 10^3 \pm 0.51$  TCID<sub>50</sub>) compared to ferrets ( $1 \times 10^6 \pm 0.70$  TCID<sub>50</sub>). Histologically, SCV infection was associated with pulmonary lesions similar to those in SCV-infected macaques, except that they were milder, especially in SCV infected cats and syncytia were not found. In the gastro-intestinal and urinary tracts SCV was detected by RT-PCR (Fig 8). Follow up of the remaining SCV-inoculated animals (n = 2 per group) revealed that they all had seroconverted by 28 d.p.i. (neutralising antibody titers 40-320). Two attempts to infect suckling mice through intracerebral inoculation failed.

Non-inoculated cats (Fig 7c, n = 2) and ferrets (Fig 7d, n = 2) housed together with the inoculated cats and ferrets, respectively, became infected with SCV; viral titers gradually increased from day 2 onwards and peaked at 6 to 8 d.p.i. Neither of the cats showed clinical signs but had seroconverted by day 28 (virus neutralising antibody titers of 40 and 160). Both ferrets showed lethargy and conjunctivitis and died on 16 and 21 d.p.i. Based on pathologic examination, the main lesions in these two animals were marked hepatic lipidosis and emaciation. There was no evidence that either of these animals died of SCV-associated pneumonia, although SCV was isolated from postmortem lung specimens of one animal.

In conclusion, domestic cats and ferrets are susceptible to experimental SCV infection and transmission of SCV to non-inoculated animals occurs efficiently. Both

species potentially could be used as animal models to test antiviral drugs or vaccine candidates against SARS.

#### 5 Example 4: SARS- interferon experiments

In a first experiment four groups of two monkeys were injected.

##### 1. PEG-INTERFERON treatment

10 Dose: 3 µg/kg or PBS injected intramuscularly according the following scheme:

Monkey:

	M001	PBS	at days -3, -1, +1 and +3
15	M002	PBS	at days -3, -1, +1 and +3
	M003	IFN	at days -3, -1, +1 and +3
	M004	IFN	at days -3, -1, +1 and +3
20	M005	PBS	at d.-3 and -1 and IFN at d.+1 and +3
	M006	PBS	at d. -3 and -1 and IFN at d. +1 and +3
	M007	IFN	at days -3, -1, +1 and +3
	M008	IFN	at days -3, -1, +1 and +3
25			

##### 2. Infection

SARS coronavirus infection of all monkeys on day 0

30

Dose: 10<sup>6</sup> TCID 50 in 5 ml PBS

- 4 ml intra-tracheal
- 1 ml intranasal
- 0.5 ml on each of the eyes

35



### 3. Sampling

a. Nose throat and rectum swabs taken on days 0, 2 and 4 and were put in 1 ml transport medium.

b. Monkeys were euthanised on day 4 and samples of lung, tracheal  
5 bronchial lymph node and trachea were harvested

Virus was cultured and titrated on Vero-118 cells, and these were scored for cytopathic affects

Virus titration using the three different swabs taken on days 0, 2 and 4 after  
infection (nose, throat and rectum) and isolation of virus from the lungs, tracheal  
10 bronchial lymph node and trachea at day 4 after infection demonstrated that the two  
control monkeys (M001 and M002) were successfully infected (table 1).

**Table 1 SARS-associated coronavirus excretion by cynomolgus  
macaques treated with pegylated interferon.**

Animal no.	specimen*					
	Pharyngeal swab			Tr. Br lymph node	Trachea	Lung
	0	2	4	4	4	4
M001	-	+	+	+	+	+
M002	-	+	+	+	+	++
M003	-	-	-	-	-	+
M004	-	-	-	+	-	+
25 M005	-	+	-	-	+	++
M006	-	+	-	+	+	++
M007	-	-	-	n.a.	n.a.	n.a.
M008	-	-	-	n.a.	n.a.	n.a.

\* day post infection

#### 1. Control animals (M001, 002)

- o Pharyngeal swabs on days 2 and 4 were all positive
- Animal M001 also was found positive with respect to isolation of SARS coronavirus from the nasal swab (day 2 and 4).
- 35 • No rectal swabs were positive

- Tissue specimen from the lungs, trachea and trachea bronchial lymph node from both control animals (M001 and M002) were positive at day 4 when the animals were sacrificed. The lung tissue homogenate contained virus at a high titer because the Vero cultures were found positive rapidly after inoculation.

5

## 2. Prophylactically treated animals (M003, 004, 007 & 008)

- negative with respect to the virus isolation test on pharyngeal swabs taken at day 0, 2 and 4 after infection (table 1).
- 10 • No nasal swab was found positive in these animals.
- Only one rectal swab of animal M004 at day 4 was scored positive (which has to be confirmed in the PCR assay because these cultures showed much bacterial contamination (cultures of rectal swabs)
- No virus isolated from trachea of M003 and 004
- 15 • Virus isolated from tracheal bronchial lymph node of M004 but not M003
- Virus isolated from lungs of M003 and 004, but are at lower titre than controls as it took longer for CPE to be observed in Vero-118 cells inoculated with samples from the lungs (confirmed by PCR – figure A below)

## 20 3. Therapeutically treated animals (M005 and 006)

### SARS coronavirus

- isolated from pharyngeal swabs taken at day 2 after infection
- not isolated from the pharyngeal swabs taken at day 4 after infection.
- 25 • isolated from more tissue samples and at higher titers from animal M005 and M006, than from animal M003 and M004 (quantitation confirmed by PCR)

Pathological examination of lung section stained by HE confirmed the low level infection of the lungs of animal M003.

30

In a second experiment treatment of *Cynomolgus* macaques was preceded by an *in vitro* dose-finding experiemnt on Vero cells.

*In vitro study*

Wells containing Vero cells were treated in triplicate with pegylated recombinant IFN- $\alpha$  (PEG-Intron, Shering Corp) for 16 h and infected with 100 TCID<sub>50</sub> per well of SCV, obtained from patient 5688, who died of SARS. After 16 h the supernatant was removed and cells were fixed by 10% neutral-buffered formalin and 70% ethanol (10 min RT). SCV antigen positive cells were visualised by immunohistochemistry, as described under histology. The number of SCV-infected cells per well was summarized as mean  $\pm$  s.d..

10 *Macaque studies.*

Three groups of cynomolgus macaques were infected intratracheally with  $1 \times 10^6$  TCID<sub>50</sub> SCV suspended in 5 ml of phosphate buffered saline (PBS). One (the control group,  $n = 4$ ) was injected intramuscularly with PBS and two (the prophylactic group,  $n = 6$  ; the post-exposure group,  $n = 4$ ) with pegylated IFN- $\alpha$  at a dose of 3  $\mu$ g/kg. The prophylactic group was injected with pegylated IFN- $\alpha$  at days -3, -1, 1 and 3 after SCV infection and the post-exposure group at days 1 and 3 after SCV infection. Four macaques from each group were euthanised at day 4 after infection. Approval for the animal experiments had been obtained from the Institutional Animal Welfare Committee. At days -3, -2, -1, 0, +2 and +4, we anaesthetised the macaques with ketamine and collected 10 ml blood from inguinal veins and took pharyngeal swabs, which were placed in 1 ml transport medium (Fouchier, R.A. *et al.*, J. Clin. Microbiol. 38, 4096-5001). Pharyngeal swabs were frozen at -70 °C until RT-PCR analysis. Pegylated IFN- $\alpha$  levels were determined using an ELISA (Bender MedSystems Diagnostics) using PEG-Intron as a standard and neopterin levels were determined as described by van Gool *et al.* (Psychiatry Res. 119, 125-132, 2003). Necropsies were done according to a standard protocol; one lung of each monkey was inflated with 10% neutral-buffered formalin by intrabronchial intubation and suspended in 10% neutral-buffered formalin overnight. Samples were collected in a standard manner (one from the cranial part of the lung, one from the medial and two from the caudal part), embedded in paraffin, cut at 5  $\mu$ m and used for immunohistochemistry (see below) or stained with haematoxylin and eosin (HE). For semiquantitative assessment of SCV-infection-associated inflammation in the lung, each HE-stained section was examined for inflammatory foci by light microscopy using a 10 $\times$  objective. Each focus was scored for size (1: smaller than or equal to area of 10 $\times$  objective, 2: larger than area of 10 $\times$  objective and smaller than or equal to area

of 2.5× objective, 3: larger than area of 2.5× objective) and severity of inflammation (1: mild, 2: moderate, 3: marked). The cumulative scores for the inflammatory foci provided the total score per animal. Sections were examined without knowledge of the identity of the macaques. The lung sections of one monkey in the post exposure group were not assessed because of the presence of inflammation from pre-mortem aspiration of food remains. Lung samples from a control group macaque were used for transmission electron microscopy as described by Kuiken, T. *et al.* (*Lancet* 362, 263-270, 2003).

Three lung tissue samples taken from the other lung (one from the cranial part of the lung, one from the medial part, and one from caudal part) were homogenised in 2 ml transport medium using Polytron PT2100 tissue grinders (Kinematica). After low speed centrifugation, the homogenates were frozen at -70° C until inoculation on Vero 118 cell cultures in 10-fold serial dilutions. The identity of the isolated virus was confirmed as SCV by RT-PCR of supernatant.

#### *Immunohistochemistry*

The same formalin-fixed paraffin-embedded lung samples as used for histology – one from the cranial part of the lung, one from the medial part, and two from caudal part – were cut at 5 µm, and stained for SCV antigen using a biotinylated purified human IgG from a convalescent SARS patient, negative control biotinylated purified human IgG, or the dilution buffer, as described by Kuiken *et al.* (*supra*). Twenty-five arbitrarily chosen 20× objective fields of lung parenchyma in each lung section were examined by light microscopy for the presence of SCV antigen expression, without knowledge of the identity of the macaques. The cumulative scores for each animal were expressed as number of positive fields per 100 fields (%). Selected lung sections from macaques in the control group were stained with anti-cytokeratin monoclonal antibody AE1/AE3 (Neomarkers) for identification of epithelial cells, according to standard immunohistochemical procedures.

#### *SCV RT-PCR*

An RT-PCR with primers and probe specific for the nucleoprotein (NP) gene of SCV was used to quantificate SCV in swabs as described by Kuiken *et al.* (*supra*). Serial dilutions of the SCV stock were used as a standard and the results were expressed as SCV eq/ml swab medium.

5

### Results

The dose finding study (3 separate experiments) on Vero cells showed a dose-dependent effect on the numbers of SCV infected cells per well. A significant effect was already observed at a dose of 1 ng/ml drug, while a dramatic reduction in the number of infected cells was observed at doses higher than 1 ng/ml (fig. 11a).

10

The control macaques showed multifocal acute DAD (diffuse alveolar damage), characterized by flooding of alveoli with protein-rich oedema fluid mixed with neutrophils and rare syncytia, extensive loss of alveolar and bronchiolar epithelium and occasional type 2 pneumocyte hyperplasia. As indicated by immunohistochemistry, there was extensive SCV antigen expression of squamous cells lining the alveolar walls. They were indicated as type 1 pneumocytes by their location, morphology, and expression of keratin in serial sections. By transmission electron microscopy, coronavirus-like particles measuring about 70 nm in diameter with typical internal nucleocapsid-like structure were found in alveolar cells. These cells were identified as type 1 pneumocytes because they lined the alveolar lumen, were closely apposed to the basement membrane, were squamous, contained abundant pinocytotic vesicles, and – in contrast to type 2 pneumocytes – had neither lamellar bodies nor microvilli. As found previously in experimentally infected macaques at 6 d.p.i., less extensive SCV antigen expression also was detected in hyperplastic type 2 pneumocytes within inflammatory foci. The combination of these histopathologic and immunohistochemical findings show that type 1 pneumocytes are the main target of SCV in early infection, and are associated with DAD.

15

20

25

30

35

High plasma levels of pegylated IFN- $\alpha$  after intramuscular injection into a group of six macaques (prophylactic group) were attained 1 day after injection (Fig. 11b), similar to peak levels found in patients after subcutaneous injection with 3  $\mu$ g/kg pegylated IFN- $\alpha$  (Bukowski, R.M. *et al.*, Cancer 95, 389-396, 2002). Because IFN- $\alpha$  is known to activate macrophages (van Gool *et al.*, *supra*), plasma levels of neopterin following pegylated IFN- $\alpha$  treatment were measured as a measure of

macrophage activation. Neopterin levels were increased in all animals (Fig. 11c), confirming the biological availability of pegylated IFN- $\alpha$  in the treated macaques.

To evaluate the prophylactic use of pegylated IFN- $\alpha$ , we experimentally infected the macaques in the prophylactic group with SCV at 3 days after the start of pegylated IFN- $\alpha$  treatment, and compared virological and pathological parameters with a control group of four macaques treated with PBS instead. We limited our investigation to the pharyngeal swabs and the lung because an earlier study did not provide evidence of extensive viral replication in other organs (Kuiper *et al.*, *supra*). We found that all parameters were significantly reduced in the prophylactic group compared to the control group. By virology, virus excretion from the pharynx was abrogated (Fig. 12), and the virus titre in the lungs at 4 d.p.i. was significantly reduced (Fig. 13a). By immunohistochemistry, the expression of SCV in type 1 pneumocytes was 90% reduced (Fig. 13b). By pathology, the extent and severity of DAD was 80% reduced (Fig. 13c). These data demonstrate that prophylactic use of pegylated IFN- $\alpha$  substantially, although not completely, protects type 1 pneumocytes of experimentally infected macaques from SCV infection, with abrogation of virus excretion and reduced severity of pulmonary lesions.

To test the efficacy of pegylated IFN- $\alpha$  as an antiviral agent post-exposure, we injected pegylated IFN- $\alpha$  intramuscularly into a post-exposure group of four macaques 1 and 3 days after experimental SCV infection, and evaluated them in the same way as the prophylactic group. Excretion of SCV from the pharynx was found only on 2 d.p.i. at a significantly reduced level compared to the control group (Fig. 12). Moreover, the virus titre in the lungs at 4 d.p.i. was significantly decreased, whereas the remaining parameters were less reduced (Fig. 13a–c). These results show that use of pegylated IFN- $\alpha$  one day post-exposure protects type 1 pneumocytes of experimentally infected macaques from SCV infection but is less effective than prophylactic use.

In this study, we have shown that type 1 pneumocytes are the main target cell for SCV infection of cynomolgus macaques early in the disease, and that pegylated IFN- $\alpha$  protects type 1 pneumocytes from SCV infection. The first point—type 1 pneumocytes as the primary target cell—is evident from the extensive presence of SCV in type 1 pneumocytes at 4 d.p.i.). The temporal sequence of lung lesions that emerges when the pathological studies in humans and macaques are viewed together is: viral infection and subsequent loss of type 1 pneumocytes; acute DAD,

characterized by flooding of alveolar lumina with highly proteinaceous oedema fluid; chronic DAD, characterized by type 2 pneumocyte hyperplasia; and, in severe cases, extensive pulmonary fibrosis. This sequence of events corresponds to the stereotypic alveolar reaction to acute lung injury from a variety of causes (Ware, L.B. and

5 Matthay, M.A., *N. Eng. J. Med.* 342, 1334-1349, 2000).

The second point—that pegylated IFN- $\alpha$  protects type 1 pneumocytes from SCV infection—is based on the beneficial effect of pegylated IFN- $\alpha$  therapy initiated 3 days before SCV inoculation of macaques. In these macaques, SCV infection of type 1 pneumocytes and severity of lung lesions were significantly reduced (Fig. 13), and  
10 viral excretion was abrogated (Fig. 12). Pegylated IFN- $\alpha$  treatment thus has an important effect on the outcome of SARS. Therefore, reduction of the viral load by pegylated IFN- $\alpha$  therapy at an early stage of SCV infection helps to prevent serious or fatal outcome of SARS associated with pulmonary fibrosis. In addition to potential disease mitigation, reduced viral excretion through pegylated IFN- $\alpha$  therapy also has  
15 an epidemiological effect by reducing the spread of SCV in the human population. Whether the mechanism of pegylated IFN- $\alpha$  protection is by direct antiviral activity or immunostimulatory effects remains to be determined.

The time interval during which effective post-exposure treatment with pegylated IFN- $\alpha$  can be initiated may be longer in humans than in the experimentally  
20 infected macaques. This is because the peak of SCV infection in the lungs is at about 16 d.p.i. in humans—based on an average incubation period of 6 days (Booth, C.M. *et al*, *JAMA* 289, 2801-2809, 2003) and a peak in viral excretion at 10 days after onset of symptoms (peiris, J.S.M. *et al*, *Lancet* 361, 1767-1772, 2003)—compared to 2 d.p.i. in these macaques (Fig. 12).

25 In conclusion, these studies show that type 1 pneumocytes are the main target cell for SCV infection of macaques early in the disease, and that pegylated IFN- $\alpha$ , a commercially available antiviral drug, protects these cells from SCV infection. Prophylactic or early post-exposure treatment with pegylated IFN- $\alpha$  will help to reduce the impact of SCV infection on healthcare workers and others possibly  
30 exposed to SCV and to limit the spread of the virus in the human population.

Claims

1. An isolated essentially mammalian positive-sense single stranded RNA virus (SARS) comprising one or more of the sequences of figure 2.  
5
2. An isolated positive-sense single stranded RNA virus (SARS) belonging to the Coronaviruses and identifiable as phylogenetically corresponding thereto by determining a nucleic acid sequence of said virus and testing it in phylogenetic tree analyses wherein maximum likelihood trees are generated using 100 bootstraps and  
10 3 jumbles and finding it to be more closely phylogenetically corresponding to a virus isolate having the sequences as depicted in figure 2 than it is corresponding to a virus isolate of BoCo (bovine coronavirus), MHV (murine hepatitis virus), AIBV (avian infectious bronchitis virus), PEDV (porcine epidemic diarrhea virus), TGEV (transmissible gastroenteritis virus) or 229E (human coronavirus 229E)..  
15
3. A virus according to claim 1 or 2 wherein said nucleic acid sequence comprises an open reading frame (ORF) encoding a viral protein of said virus.
4. A virus according to claim 3 wherein said open reading frame is selected from  
20 the group of ORFs encoding the viral replicase, nuclear capsid protein and the spike protein.
5. A virus according to claim 1-4 isolatable from a human with atypical pneumonia.  
25
6. An isolated or recombinant nucleic acid or SARS virus-specific functional fragment thereof obtainable from a virus according to anyone of claims 1 to 5.
7. A vector comprising a nucleic acid according to claim 6.  
30
8. A host cell comprising a nucleic acid according to claim 6 or a vector according to claim 7.



9. An isolated or recombinant proteinaceous molecule or SARS virus-specific functional fragment thereof encoded by a nucleic acid according to claim 6.
10. An antigen comprising a proteinaceous molecule or SARS virus-specific functional fragment thereof according to claim 9.
11. An antibody specifically directed against an antigen according to claim 10.
12. A method for identifying a viral isolate as a SARS virus comprising reacting said viral isolate or a component thereof with an antibody according to claim 11.
13. A method for identifying a viral isolate as a SARS virus comprising reacting said viral isolate or a component thereof with a nucleic acid according to claim 6.
14. A method for virologically diagnosing a SARS infection of a mammal comprising determining in a sample of said mammal the presence of a viral isolate or component thereof by reacting said sample with a nucleic acid according to claim 6 or an antibody according to claim 11.
15. A method for serologically diagnosing a SARS infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a SARS virus or component thereof by reacting said sample with a proteinaceous molecule or fragment thereof according to claim 9 or an antigen according to claim 10.
16. A diagnostic kit for diagnosing a SARS infection comprising a virus according to anyone of claims 1 to 5, a nucleic acid according to claim 6, a proteinaceous molecule or fragment thereof according to claim 9, an antigen according to claim 10 and/or an antibody according to claim 11.
17. Use of a virus according to any one claims 1 to 5, a nucleic acid according to claim 6, a vector according to claim 7, a host cell according to claim 8, a proteinaceous molecule or fragment thereof according to claim 9, an antigen according to claim 10,

or an antibody according to claim 11 for the production of a pharmaceutical composition.

18. Use according to claim 17 for the production of a pharmaceutical composition  
5 for the treatment or prevention of a SARS virus infection.

19. Use according to claim 17 or 18 for the production of a pharmaceutical composition for the treatment or prevention of atypical pneumonia.

10 20. A pharmaceutical composition comprising a virus according to any one of claims 1 to 5, a nucleic acid according to claim 6, a vector according to claim 7, a host cell according to claim 8, a proteinaceous molecule or fragment thereof according to claim 9, an antigen according to claim 10, or an antibody according to claim 11.

15 21. A method for the treatment or prevention of a SARS virus infection comprising providing an individual with a pharmaceutical composition according to claim 20.

22. A method for the treatment or prevention of atypical pneumonia comprising providing an individual with a pharmaceutical composition according to claim 20.

20

23. A viral replicase encoded by an RNA sequence comprising the sequences EMC-1, EMC-2, EMC-3, EMC-4, EMC-5, EMC-6, EMC-7, EMC-13 and/or EMC-14, or homologues thereof as depicted in figure 2.

25 24. A viral spike protein comprising the amino acid depicted as translation 2 with sequence EMC-7 and translation 1 of RDG 1 as depicted in figure 2, or a homologue thereof.

25 A viral nuclear capsid protein encoded by an RNA sequence comprising the  
30 sequence EMC-8 as depicted in figure 2 or a homologue thereof.

26. A viral protein encoded by an RNA sequence comprising the sequence EMC-9, EMC-11 and/or EMC-12 as depicted in figure 2.

27. A nucleic acid sequence which comprises one or more of the sequences EMC-1 to EMC-13 as depicted in figure 13 or a nucleic acid sequence which can hybridise with any of these sequences under stringent conditions.
- 5 28. Use of interferon for the preparation of a medicament for the treatment or prevention of a coronavirus associated disease.
29. Use according to claim 28 wherein said interferon is interferon alpha.
- 10 30. Use according to claim 29, wherein said interferon is interferon-alpha 2a.
31. Use according to claim 29, wherein said interferon is interferon-alpha 2b.
- 15 32. Use according to any of claims 28-31, wherein said interferon is pegylated.
33. Use according to any of claims 28-32, wherein said coronavirus associated disease is a disease of animals, preferably vertebrates, more preferably birds or mammals, especially humans, ape or rodent.
- 20 34. Use according to claim 33, wherein said disease is a respiratory disease and/or gastroenteritis.
- 25 35. Use according to claim 33 or claim 34, wherein said animal is human.
36. Use according to any of claims 28-35 wherein said coronavirus associated disease is a disease caused by HCoV-NL, the feline infectious peritonitis virus (FIPV) or hemagglutinating encephalomyelitis virus (HEV) of swine or avian infectious bronchitis virus (IBV) or mouse hepatitis virus (MHV).
- 30 37. Use according to any of claims 28-35 wherein said coronavirus associated disease is a disease caused by a SARS coronavirus (SARS-CoV).

38. Use according to claim 37, wherein said SARS virus is a positive-sense single stranded RNA virus (SARS coronavirus) comprising one or more of the sequences of figure 2.

5

39. Use according to claim 28, wherein said SARS virus is a positive-sense single stranded RNA virus (SARS coronavirus) corresponding to GenBank accession no. AY274119 or AY278741 or AY338175 or AY338174 or AY322199 or AY 322198 or AY322197 or AH013000 or AY322208 or AY322207 or AY 322206 or AY322205 or  
10 AH012999.

40. A method for the treatment or prevention of a coronavirus associated disease in an animal, preferably a vertebrate, more preferably a bird or mammal, especially human, ape or rodent, infected with a coronavirus, said method comprising  
15 administering interferon, to said animal, preferably a vertebrate, more preferably a bird or mammal, especially human, ape or rodent, along with a pharmaceutically acceptable carrier.

41. A method according to any of claims 40 wherein said interferon is  
20 administered together with a vaccine, antibody and/or antiviral agent.

42. A method according to claim 41, wherein said vaccine, antibody and/or anti-viral agent is selected from the group consisting of whole inactivated virus vaccines, attenuated vaccines, sub-unit vaccines, recombinant vaccines, antibody for passive  
25 immunization, nucleoside analogs such as ribavirin, RNA-dependent RNA polymerase inhibitors, protease inhibitors.

30

Figure 1.

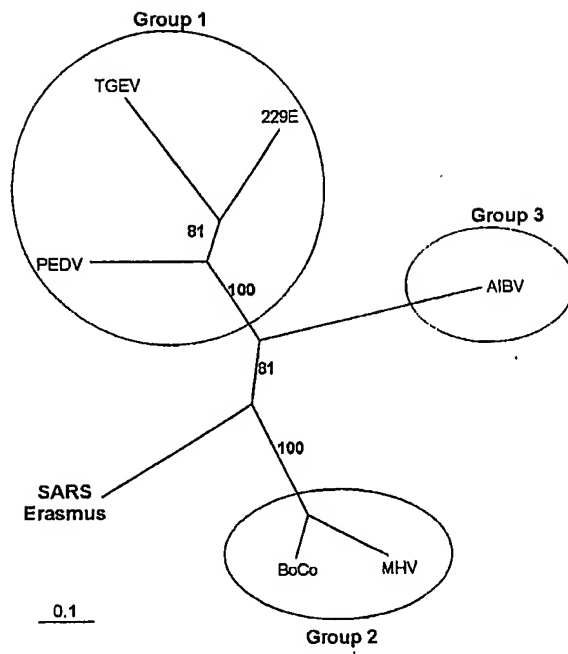


Figure 2 RNA sequences, implied polypeptides and alignment with one close relative

# **EMC-1**

5 UUGUAAACUGGUGGUCUUGUACAACAGACUUCUCAGUGGUUGUCUAAUCUUUUGGGCACUACUGGUUGAAAAAC  
UCAGGCCUAUCUUUGAAUGGAUUGAGGCGAAACUUAGUGCAGGAGUUGAAUUUCUAAGGAUGCUUGGGAGAU  
UCUCAAAUUUCUCAUUAACAGGUGUUUUUGACAUCGUCUAGGGUCAAUACAGGUUGCUUCAGAUAAACAUAAG  
GAUUGUGUAAAAUGCUUCAUUGAUGUUGUUAACAGGCACUCGAAAUGUGCAUUGAUAAGUCACUAUCGCUG  
10 GCGCAAAGUUGCGAUCACUACAACUAGGUGAAGUCUUCUACGCUCUAAAGCAAGGGACUUUACCGUCAGUGUAU  
ACGUGGCAAGGAGCAGCUGCAACUACUAGCCUCUUAAGGCACCAAAAGAAGUAACCUUUCU  
UGAAGGUGAUUCACAUGACACAGUACUACCUCUGAGGAGGUUGUUCUCAAGAACGGUGAA  
CUCGAAGCACUCGAGACGCCCUGUAGCUUCACAAUUGGAGCUAUCGUUGGCACACCAG  
UCUGUGUAAAUGGCCUCAUGCUCUAGAGAUUAAGGACAAAGAACAUAUCUGCGCAUUGUC  
15 UCCUGGUUUACUGGCUACAACAACUUGUCUUUCGCUUAAAAGGGGGUGCACCAAUUAAAGGU  
GUAACCUUUGGAGAAGAUACUGUUUGGGAAGUUCAGGGUUAACAAGAUUGUGAGAAUCACAU  
UUGAGCUUGAUGAACGUGUUGACAAAGUGCUUAAUGAAAAGUGCUCUGUCUACACUGUUGA  
AUCCGGUACCGAAGUACUGAGUUUGCAUGUGUUGUAGCAGAGGCUGUUGUGAAGACUUUA  
CAACCAGUUUCUGAUC

## 20 **Translation**Nucleotides 7 to 870: Frame 1; 288 aa

LVLVLYNRLLSGLIFWALLVEKLRPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV  
NKALEMCIDQVTIAGAKLRSLNLGEVFIQAQSKGLYRQCIRGKEQLQLLMPKAPKEVTFLEGDSHDTVLTSEEVVLKNGEL  
EALETVPVDSFTNGAIVGTPVCVNGLMLEIKDKEQYCALSPGLLATNNVFRLLKGGAPIKGVTFGEDTVWEVQGYKNVRITF  
ELDERVDKVLNEKCSVYTVESGTEVTEFACVVAEAVVKTLPVSD

## 25 **Alignment**

RNA-directed RNA polymerase (orfla) murine hepatitis virus  
Identities = 72/285 (25%), Positives = 118/285 (41%)

30 Query: 49 FWALLVEKLRPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCV 228  
F AL V +R I EW + L+ + W + L+ G+F + G I + + + V  
Sbjct: 638 FKALGVAVVRKITEWFD--LAVDIAASAAGWLCYQ-LVNGLFAVANGVITFVQE-VPSELV 693

35 Query: 229 KCFIDVVNKALEMCIDQVTIA---GAKLRSLNLGEVFIQAQSKGLYRQCIRGKEQLQLLMP 399  
K F+D ++ ID ++++ G + V +A SK +Y + K +MP  
Sbjct: 694 KNFVDKFAFFKVLIDSMVSILSGLTVVKTASNRVCLAGSK-VYE--VVQKLSAYVMP 750

40 Query: 400 LKAPKEVTFLEGDSHDTVLTSEEVVLKNGEL--EALETVPVDSFTNGAIVGTPVCVNGLM 573  
+ E T L G+ V + V + L + P SF IV L  
Sbjct: 751 VGC-SEATCLVGEIEPAVFEDDVVDVVKAPLTYQGCKPPTSFEKICIVDK-----L 801

45 Query: 574 LEIKDKEQYCAL-----SPGLLATNNVFRLLKGGAPIKGVTFGEDT-VWEVQGYKNVRITF 735  
K +Q+ + + G+L F G K V F + V ++ + ++ITF  
Sbjct: 802 YMAKCGDQFYFPVVVDNDTVGVLDQCWRFPACAG---KKVEFNDKPKVRKIPSTRKIKITF 857

50 Query: 736 ELDERVDKVLNEKCSVYTVESGTEVTEFACVVAEAVVKTLPVSD 870  
LD D VL++ CS + V+ + E VV +AV TL P +  
Sbjct: 858 ALDATFDSVLSKACSEFEVDKDVTLDELDDVVLDAVESTLSPCKE 902

# **EMC-14**

CAUCCAGCUUCUUAAGGCAGCAUAUGAAAAUUUCAAUUCACAGGACAUCUUAUUGCACCAUUGUUGUCAGCA  
GGCAUUAUUGGUGCUAAACCACUUCAGUCUUUACAAGUGUGCGUGCAGACGGUUCGUACACAGGUUUUAUUG  
CAGUCAUUGACAAAGCUCUUAUAGAGCAGGUUGUCAUGGAUUAUCUUGAUAACCUAGAGCCUAGAGUGGAAGC  
55 ACCUAAACAAGAGGAGCCACCAACACAGAAGAUUCCAAAACUGAGGAGAAAUCUGUCGUACAGAAGCCUGUC  
GAUGUGAAGCCAAAAAUUAGGCCUGCAUUGAUGAGGUUACCACAACACUGGAAGAAACUAAGUUUCUUAACA  
AUAAGUUACUCUUGUUUGCUGAUUAUUAUGGUAAGCUUUAACCAUGAUUCUCAGAACAUUGCUUAGAGGUGAAGA  
UAUGUCUUUCCUUGAGAAGGAUGCACCUCUUAUGGUAGGUGAUGUUAUCACUAGUGGUGAUAUCACUUGUGUU

Fig. 2 Cont.

GUAUAUACCCUCCAAAAAGGCUGGUGGCACUACUGAGAUGCUCUCAAGAGCUUUGAAGAAAGUGCCAGUUGAUG  
 AGUAUAUAACACGUAACCCUGGACAAGGAUGUGCUGGUUAUACACUUGAGGAAGCUAAGACUGCUCUUAAGAA  
 AUGCAAUUCUGCAUUUUAUGUACUACCUUCAGAAGCACCUAAGCUAAGGAAGAGAUUCUAGGAACUGUAUCC  
 UGGAAUUGAG

5

*Translation*

Nucleotides 5 to 739: Frame 2; 245 aa

10

IQLLKAAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQTVRTQVYIAVNDKALYEQVVM DYLDNLKPRVEAPKQEEPPN  
 TEDSKTEEKSVVQKPV DVKPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYMGDVI  
 TSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTPGQGCAGYTL EEAKTALKKCKSAFYVLPSEAPNAKEEILGTVS  
 WN

*Alignment*

15

replicase polyprotein lab Human coronavirus 229E

Identities = 48/202 (23%), Positives = 83/202 (41%), Gaps = 13/202 (6%)  
 Frame = +2

20

Query: 8 LLKAAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQTVRT---QVYIAVNDKALYEQV 178  
 L+KA N Q L P+LS GIFG K SL+V + T +V++ + + +  
 Sbjct: 1371 LIKAYNTINNEQGTPLTPILSCGIFGIKLETSLEVLLDVCNTKEVKVFVYTDTEVCKVKD 1430

25

Query: 179 VMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPV DVKPKIKACIDEVTTTLEETKF 358  
 + L N++ +VE PK E P V KP V K +++ ++  
 Sbjct: 1431 FVSGLVNVQ-KVEQPKIEPKP-----VSVIKVAPKPYRVDGKFSYFTEDLLCVADDKPI 1483

30

Query: 359 L--TNKLLLFADINGKLYHDSQNMLRG--EDMSFLEKDAP-----YMGDVI TSGDITC 508  
 + T+ +L D L + +L +D + K P + +G V+ +  
 Sbjct: 1484 VLFTDSMLTLDDRGLALDNALSGVLSAAIKDCVDINKAIPSGNLIKFDIGSVV----VYM 1539

Query: 509 VVIPSKKAGGTTEMLSRALKKV 574  
 V+PS+K + R +K+  
 Sbjct: 1540 CVVPSEKDKHLDNNVQRCTRKL 1561

35

**EMC-2**

UCGAGAUUUcAUcUUGACGGUGCAGGUUCUUCACUUGACAAACUAAAGAGUCUCUUAUCCUGCGGGAGGUU  
 AAGACUAUAAAAGUGUUCACAACUGUGGACAACACUAAUCCACACACAGCUUGUGGAUAUGUCUAUGACAU  
 AUGGACAGCAGUUUGGUCCAACAUACUUGGAUGGUGCUGAUGUUACAAAAUUAACCUCUAUGUAAAUCAUGA  
 GGGUAAGACUUUCUUGUACUACCUAGUGAUGACACACUACGUAGUGAAGCUUUCGAGUACUACCAUACUCUU  
 GAUGAGAGUUUUCUUGGUAGGUACAUGUCUGCUUUAACACACAAAGAA AUGGAAA

40

*Translation*

Nucleotide 2 to 349: Frame 2; 116 aa

45

RDFILTVQVLSLDKLSLLSLREVKTIKVFTTV DNTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTF FVLP  
 SDDTLRSEAF EYYHTLDESFLGRYMSALNHTKKWK

50

*Alignment*

> Bovine Coronavirus RNA-Dependent RNA polymerase

Identities = 25/90 (27%), Positives = 44/90 (48%)  
 Frame = +2

55

Query: 80 IKVFTTV DNTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTF FVLP SDDTL 259  
 + + TVD N + V + ++G+ G + DG +VTK K +N++GK FF + +  
 Sbjct: 1565 VDILLTV DGVNFTNRFPVGE SFGKSLGNVFC DGVNVT KHKCDIN YKGKVFFQFDNLSSE 1624

60

Query: 260 RSEAF EYYHTLDESFLGRYMSALNHTKKWK 349  
 +A D+ L Y + L + KW+  
 Sbjct: 1625 DLKAVRSSFNF DQKELLAYYNMLVNC SKWQ 1654

65

Fig. 2 Cont.

**EMC13:**

CUGAAGAAGUAGUGGaAAAUCCUACCAUACAGAAGGAAGUCAUAGAGUGUGACGUGAAAAACUACCGAAGUUGU  
 AGGCAAUGUCAUACUAAACCAUCAGAUGAAGGUGUAAAGUAAACACAAGAGUUAGGUCAUGAGGAUCUUUUG  
 GCUGCUUUGUGGAAAACACAAGCAUUAACCAUUAAGAAACCUAAUGAGCUUUCACUAGCCUUAGGUUUAAAAA  
 5 CAUUGCCACUCAUGGUUUGCUGCAUUAUAGUGUCCUUGGAGUAAAAUUUGGCUUAGUCAACCAU  
 CUUAGGACAAGCAGCAAUUACAACAUCAAAUUGCGCUAAGAGAUUAGCACAACGUGUGUUUAAACAAUUAUUG  
 CCUUUUGUGUUUACAUUAUUGUCCAAUUGUGUACUUUUUACUAAAAGUACCAUUCUAGAAUUAGAGCUUCAC  
 UACCUACAACUUAUUGCUAAAAUAGUGUUAAGAGUGUUGCUAAAAUUAUGUUUGGAUGCCGGCAUUAUUUUGU  
 GAAGUCACCCAAAUUUUCUAAAUUGUUCACAAUCGCUAUGUGGCUAUGUUGUUAAGUAAUUGCUUAGGUUCU  
 10 CUAAUCUGUGUAAUCUGCUGCUUUUGGUGUACUCUUAUCUAAUUUUGGUGCUCCUUCUUAUUGUAAUGGCGUUA  
 GAGAAUUGUAUCUUAUUCGUCUAAAGUUAUACUUAUGGAUUUCUGUGAAGGUUCUUUCCUUGCAGCAUUUG  
 UUUAAAGUGGAUUAAGACUCCUUGAUUCUUAUCCAGCUCUUGAAACCAUUCAGGUGACGAUUUCAUGGUACAAG  
 CUAGACUUGACAAUUUUGGUCUGGCCGUG

15 **Translation**

>-out: 3 to 833; Frame 3 277 aa  
 EEVVENPTIQKEVIECDVKTTTEVVGNVILKPSDEGVKVTQELGHEDLMAAYVENTSITIKKPNELSLALGLKTIATHGIAA  
 INSVFWSKILAYVKPFLGQAAITTSNCAKRLAQRVFNMYVFTLLFQLCTFTKSTNSRIRASLPTTIKNSVKSVAKLC  
 LDAGINYVKSFKFLFTIAMWLLLSICLGLSICVTAAFGVLLSNFGAPSYCNGVRELYLNSSNVTMTDFCEGSFPCSI  
 20 LSGLDSDLSYPALETITQVTISSYKLDLTILGLAA

**Alignment**

bovine coronavirus RNA-dependent RNA Polymerase

Identities = 50/269 (18%),

25 Query: 57 KTTTEVVGNVILKPSDEGVKVTQELGHEDLMAAYVENTSITIKKPNELSLALGLKTIATH- 233  
 K +V +VI+ +K + L D+ ++ ++ N+LS+A+ + TI  
 Sbjct: 2046 KPFKVEDSVIVNDDTSEIKYKSLSIVDVYDMWLTGCRYVVRTANDLSMAVNVPTIRKFI 2105  
 30 Query: 234 --GIAAINSVPWSKI-LAYVKPFLGQAAITTSNCAKRLAQRVFN--NYMPYVFTLLF--- 389  
 G+ + S+P + L +KP N K + ++ N++ ++F LLF  
 Sbjct: 2106 KFGMTLV-SIPIDLNLREIKPVF-----NVVKAVRNKISACFNFIKWLFLVLLFGWI 2156  
 35 Query: 390 -----QLCTFTKSTNSRIRASLPTTIKNSVKSVAKLCLDAGINYVKSPKFSKLFTIAMW 554  
 +T S++ L KN+ + + G + + +W  
 Sbjct: 2157 KISADNKVIYTTTEVASKLTCKLVALAFKNAFLTFKWSVVARGACIIAT-----IFLLW 2209  
 Query: 555 XXXXXXXXXXXXXVTAAFGVLLSNFGAPSYCNGVRELYLNSSNVTM----- 695  
 G L P++ + + ++ ++ T+  
 40 Sbjct: 2210 FNFIYANVIFSDFYLPKIGFL-----PTFVGKIAQWIKSTFSLVTICDLYSIQDVGFKN 2263  
 Query: 696 DFCEGSFPCSIKLSGLDSDLSYPALETIQ 782  
 +C GS C CL+G D LD+Y A++ +Q  
 Sbjct: 2264 QYCNGSIACQFCLAGFDMLDNYKAIDVVQ 2292

45

**EMC-3**

GUGGUAAGAUUGUAGUACUUGUUUAAACUUAUGCUUAAGGCCACAUAUUGUGCGUUCU  
 UGCUGCAUUAAGUUUGUUAUUAUCGUUAUGCCAGUACAUAUUGUCAUCCAUGAUGGUUAC  
 ACAAAUGAAAUAUUGGUUACAAAGCCAUUCAGGAUGGUGUCACUCGUGACAUCAUUCUA  
 50 CUGAUGAUUGUUUGCAAUAUAAACAUUGCUGGUUUUGACGCAUGGUUUAGCCAGCGUGGUGG  
 UUCAUACAAAAAUGACAAAAGCUGCCUGUAGUAGCUGCUAUCUUAACAAGAGAGAUUGGU  
 UUCAUAGUGCCUGGCUUACCGGGUACUGUGCUGAGAGCAAUCAUUGGUGACUUCUUGCAU  
 UCCUACCUCGUGUUUUAGUGCUGUUGGCAACAUUGCUACACACCUUCCAAACUCAUUGA  
 GUUAUGUGAUUUUGCUACCUCU

55

**Translation**

Nucleotide 3-449; 149 aa

GKIVSTCFKMLKATLLCVLAALVCYIVMPVHTLSIHGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFD  
 AWFSQRGGSYKNDKSCPVVAIITREIGFIVPGLPGLTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDF  
 60 ATS



## Fig. 2. Cont.

## Alignment

&gt; Murine Hepatitis Virus RNA-Dependent RNA polymerase

5 Identities = 48/126 (38%),

Query: 78 YIVMPVHTLSIHGTYTNEIIGYKAIQDGVTRDIIISTDDCFANKHAGFDAWFSQRRG--SY 251  
 + +MP + + D +K I +GV RD+ TD CFANK FD W+ G Y  
 10 Sbjct: 2859 WALMPTYAVHKSDMQLPLYASFVKVIDNGVLRDVSVDACFANKFNQFDQWYESTFGLAYY 2918

Query: 252 KNDKSCPVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEY 431  
 +N K+CPVV A+I ++IG + +P TVLR LHF+ F+ CYTP I Y  
 Sbjct: 2919 RNSKACPVVVAVIDQDIGHTLFNVPTTVLR-YGFHVLHFITHAFATDSVQCYTPHMQIPY 2977

15 Query: 432 SDFATS 449  
 +F S  
 Sbjct: 2978 DNFYAS 2983

EMC-4

20 ACAGACAUCAAUCACUUCUGCUGUUCUGCAGAGUGGUUUUAGGAAAAUGGCAUCCCCGUCAGGCAAAGUUGAA  
 GGGUGCAUGGUACAAGUAACCUGUGGAACUACAACUCUUAUUGGAUUGUGGUUGGAGACACAGUAUACUGUC  
 CAAGACAUGUCAUUGCAGCAGACAAGACAUGCUUAAUCCUAACUAUGAAGAUUCGCUCAUUCGCAAAUCCAA  
 CCAUAGCUUUUCUUGUUCAGGCUGGCAUGUUAACUUCGUGUUAUUGGCCAUUCUAUGCAAAAUUGUCUGCUU  
 AGGCUUAAAAGUUGAUACUUCUAACCCUAAGACACCCAAGUAUAAAUUGUCCGUAUCCAACCGUGUCAACAU  
 25 UUCAGUUCUAGCAUGCUACAAGGUUACCAUCUGGUGUUAUACAGUGGCCAUGAGACCUAAUCAUACCAU  
 UAAAGGUUCUUCUUAUUGGAUCAUGUGGUAGUGUUGGUUUUAACAUUGAUUAUGAUUGCGUGUCUUCUGC  
 UUAUGCAUCAUAUGGAGCUUCCAACAGGAGUACACGCUGGUACUGACUUAAGAAGGUAAAUCUAUGGUCCAU  
 UUGUUGACAGACAAACUGCAGGCGUACAGACACAACCAUAACAUAUAAUGUUUUGGCAUGGCUGUA  
 UGCUGCUGUUAUCAUGGUGAUA

30

## Translation

Nucleotides 2 to 679: Frame 2; 226 aa

35 QTSITSAVLQSGFRKMAFPPSGKVEGCMVQVTCGTTTLNGLWLDLDTVYCPRHVICTAEDMLNPYEDLLIRKSNHSLVQAG  
 NVQLRVIGHSMQNCLRLKVDTSNPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMPNHTIKGSFLNGSCGSGVFNI  
 DYDCVSFCYMHMELPTGVHAGTDLEGKFYGFVDRQTAQAAGTDTTITLNLAWLYAAVINGD

## Alignment

RNA-directed RNA polymerase murine hepatitis virus

40

Identities = 122/222 (54%)

Query: 8 SITS AVLQSGFRKMAFPPSGKVEGCMVQVTCGTTTLNGLWLDLDTVYCPRHVICTAEDMLNP 187  
 S+T++ LQSG KM P+ KVE C+V VT G TLNGLWLDLDTVYCPRHVIC++ DM +P  
 45 Sbjct: 3326 SVTTSFLQSGIVKMSVPTSKVEPCIVSVTYGNMTLNLGLWLDLDTVYCPRHVICSSADMTDP 3385

Query: 188 NYEDLLIRKSNHSLVQAGNVQLRVIGHSMQNCLRLKVDTSNPKTPKYKFVRIQPGQTF 367  
 +Y +LL R ++ F V +G + L V+ + MQ C L L V NP TP KY F ++PG+TF  
 50 Sbjct: 3386 DYPNLCRVTSDDFCVMSGRMSLTVMYQMQGCQLVLTVTLQNPNTPKYSFGVVKPGETF 3445

Query: 368 SVLACYNGSPSGVYQCAMPNHTIKGSFLNGSCGSGVFNI DYDCVSFCYMHMELPTGVH 547  
 +VLA YNG P G + +R +HTIKGSFL GSCGSGV+ + D V F YMH +EL TG H  
 Sbjct: 3446 TVLAAAYNGRPQGAHFVTLRSSHTIKGSFLCGSCGSGVGYVLTGDSVRFVYMHQLELSTGCH 3505

55 Query: 548 AGTDLEGKFYGFVDRQTAQAAGTDTTITLNLAWLYAAVIN 673  
 GTD G FYGP+ D Q Q D T T+NV+AWLYAA+ N  
 Sbjct: 3506 TGTDFSGNFYGPYRDAQVVQLPVQDYTTQTVNVVAVLYAAIFN 3547

EMC-5

60 Note that this sequence is not fully in frame.

AGUUGGAAAAGAUGGCAGAUACAGGCUAUGACCCAAAUGUACAAACAGGCAAGAUCUGAGGA  
 CAAGAGGGCAAAAAGUAAACUAGUGCUAUGCAAACAAUGCUCUUCACUAUGCUUAGGAAGCUU  
 GAUAAUGAUGCACUUAACAACAUAUAACAACAAUGCGCGUGAUGGUUGUGUCCACUCAACA  
 UCAUACCAUUGACUACAGCAGCCAAACUAUGGUUGUUGUCCUGAUUAUGGUACCUACAA  
 65 GAACACUUGUGAUGGUAAACACCUUUAACAUAUGCAUCUGCACUCUGGGAAAUCCAGCAAGUU  
 GUUGAUGCGGAUAGCAAGAUUGUUAACUUAUGUAAAUAACAUGGACAAUUCACCAAAU  
 UGGCUUGGCCCUUAUUGUUAACAGCUCUAAGAGCCAAACUCAGCUGUUAACUACAGAAUAA  
 UGAACUGAGUCCAGUAGCACUACGACAGAUGUCCUGUGCGGCUGGUACCACACAAACAGCU  
 UGUACUGAUGACAAUGCACUUGCCUACUAUAACAUAUUCGAAGGGAGGUAGGUUUGUGCUGG

Fig. 2. Cont.

CAUUACUAUCAGACCACCAAGAUCUCAAAUGGGCUAGAUUCCCUAAGAGUGAUGGUACAGG  
 UACAAUUUACACAGAACUGGAACCACCUUGUAGGUUUGUUACAGACACACCAAAGGGCCU  
 AAAGUGAAAUACUUGUACUUAUCAAAGGCUUAAACAACCUAAAUAGAGGUAUGGUGCUGGG  
 CAGUUUAGCUGCUACAGUACGUCUUCAGGCUGGAAAUAGCUACAGAAGUaCCUGCCAAUUA  
 5 ACUGUGCUUUCUUCUGUGCUUUUGCAGUAGACCCUGCUAAAAGCAUAUaAAGGAUUACCUA  
 GCAAGUGGAGGACAACCAAUACCAACUGUGUGAAGAUGUUGUGUACACACACUGGUACAG  
 GACAGGCAAUUACUGUAACACCAGAAGCUAACAUGGACCAAGAGUCCUYUGGUGGUCUUC  
 AUGUUGUCUGUAUUGUAGAUGCCACAUGACCAUCCAAAUCCUAAAGGAYUCUGUGACUUG  
 AAAGGUAAGUACGUCCAAAUACCACUUGUGCUAAUGACCCAGUGGGUUUACACUUA  
 10 GAAACACAGUCUGUACCGUCUGCGGAUUGUGGAAAGGUUAUGGCUGUAGUUGUGACCAACU  
 CCGCGAACCCUUGAUGCAGUCUGCGGAUGCAUCAMCGUUUUUAAACGGGUUUGCGGUGUAA  
 GUGCAGCCCUCUACACCGUGCGGCACAGGCACUAGUACUGAUGUCGUCUACAGGGCUUU  
 UGAUAUUUACAACGAAAAGUUGCUGGUUYUGCAAAGUCCUAAAAACUAA

15 *Translation 1*

Nucleotide 3-701 ; 233 aa

LEKMADQAMTQMYKQARSEDKRAKVTSAMQTMFTMLRKLDNDALNNIINNARDGCVPLNIPLTTAAKLMVV  
 VPDYGTYNKTCGNTFTYASALWEIQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANSVKLQNNELSP  
 VALRQMSCAAGTTQTACTDDNALAYYNNKGGRFVLALLSDHQLKWARFPKSDGTGTIYTELEPPCRFVTD  
 20 PKGPKVKYLYFIKA

*Translation 2*

FKRVCGVSA-ARLTPCGTGTSTDVVYRAFDIYNEKVAGXAKFLK

25 *Alignment 1 of translation 1 sequence*

RNA-Dependent RNA Polymerase: bovine coronavirus  
 Identities = 181/413 (43%),

Query: 3 LEKMADQAMTQMYKQARSEDKRAKVTSAMQTMFTMLRKXXXXXXXXXXXXXRDGCVPLN 182  
 LE+MAD A+T MYK+AR DK++KV SA+QTMFLF+M+RK GCVPLN  
 30 Sbjct: 3985 LERMADLALTNMKEARINDKSKVVSALQTMFLSMVRKLDNQALNSILDNAVKGCVPLN 4044

Query: 183 IIPLTAAKLMVVVPDYGTYNKTCGNTFTYASALWEIQVVDADSKIVQLSEINMDNSP 362  
 IP A L ++VPD Y D TYA +W+IQ + D+D QL+EI+ D +  
 35 Sbjct: 4045 AIPSLAANTLTIIIPDKSVYDQVVDNVYVYAGNVWQIQTIQSDGTNKQLNEISDDCN- 4103

Query: 363 NLAWPLIVTALRAN--SAVKLQNNELSPVALRQMSCAAGTTQTACTDDNALAYYNNKGG 536  
 WPL++ A R N SA LQNNEL P L+ +G QT T YNNNS G  
 40 Sbjct: 4104 ---WPLVLIANRHNEVSATVLQNNELMPAKLKTQVNSGPDQTCNTPTQ--CYNNNSNG 4158

Query: 537 RFVLALLSDHQLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPKVKYLYFIKA\*TT\*I 716  
 + V A+LSD LK+ + K DG + EL+PPC+F KG K+KYLYF+K T  
 Sbjct: 4159 KIVYAILSDVDGLKYTKILKDDG-NFVVLELDPPCKFTVQDVKGLKIKYLYFVKGCNTLA 4217

45 Query: 717 EVWCNAV\*LLQYVFR-----EMLQKYLPIQLCFPSVLLQ\*TLKHKIDYLASGGQPI 878  
 W V + RL E + LC SV + T L D++ GG PI  
 Sbjct: 4218 R--GWVGTISSTVRLQAGTATEYASNSILSLCAFSVDPKKTLYL----DFIQGGTPIA 4271

50 Query: 879 NCVKMLCTHTGTGQAITVTPEANMDQESXGGASCCLYCRCHIDHPNPKGXCDLKGKYVQI 1058  
 NCVKMLC H GTG AITV P+A +Q+S GGAS C+YCR ++HP+ G C L+GK+VQ+  
 Sbjct: 4272 NCVKMLCDHAGTGMAITVKPDATTNQDSYGGASVCIYCRARVEHPDVGDLCKLRGKFVQV 4331

Query: 1059 PTTCANDEPVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQSADASXFLNGFAV 1217  
 P DPV + L + VC VCG W+ CSC + +QS D + FLNGF V  
 55 Sbjct: 4332 PVG-IKDPVSYVLTHDVCQVCGFWRDGCSCVS-TDTTVQSKDTN-FLNGFGV 4381

*Alignment 2 of translation 2 sequence*

RNA-directed RNA polymerase (ORF1B) [murine hepatitis virus]

Identities = 24/44 (54%),

60 Query: 1199 FKRVCGVSA-ARLTPCGTGTSTDVVYRAFDIYNEKVAGXAKFLK 1327  
 FKRV G S ARL PC +G TDV RAFDI N AG + K  
 65 Sbjct: 18 FKRVRGTSVNARLVPCASGLDQVQLRAFDICNANRAGIGLYK 61



Fig. 2. Cont.

GAUUAUUUAACAAUUCUACUAAUGUUGUUUAUACGAGCAUGUAAACUUUGAAUUGUGUGACAACCCUUUCUUU  
 GCUGUUUCUAAACCAUGGGUACACAGACACAUACUAGUAAUUCGAUAAUGCAUUUAAUUGCACUUUCGAGU  
 ACAUAUCUGAUGCCUUUCGCUUGAUUUUCAGAAAAGUCAGGUAUUUUUAAACACUUACGAGAGUUUGUGUU  
 UAAAAUAAAGAUGGGUUUCUUAUGUUUAUAGGGCUAUCACCUAUAAGAUAGUUGCUGAUUCUACCUUCU  
 5 GGUUUUAACACUUUGAAACCUAUUUUUUAGUUGCCUCUUGGUUUUAACAUUACAAUUUUUAGAGCCAUUCUUA  
 CAGCCUUUCACCUGCUCAAGACAUUUGGGGCACGUCAGCUGCAGCCUAAUUUUUUGGCUAUUUUAAAGCCAAC  
 UACAUUUUAGCUCUAGUAUGAUGAAAUGGUACAUCACAGAUGCUGUUGAUUGUUCUAAAUCCACUUGCU  
 GAACUCAAAAGCUCUGUUUAGAGCUUUGAGAUUGACAAGGAAUUUACCAGACCUCUAAUUUCAGGGUUGUUC  
 CCUCAGGAGAUGUUGGAGAUUCCCUAAUUAUUAACAAUUGUUGCCUUUUUGGAGAGGUUUUUAAUGCUACUAA  
 10 AUUCCCUUCUGUCUUAUGCAUGGGAGAGAAAAAAAUUUCUAAUUGUGUUGCUGAUUACUCUGUGCUCUACAAC  
 UCAACAUUUUUUUAACCUUUUAAGUGCUAUGGCGUUCUGCCACUAAGUUGAAUGAUUUUGCUUCUCCAUG  
 UCUAUGCAGAUUCUUUUGUAGUCAAGGGAGAUGAUGUAAGACAAUAGCGCCAGGACAAACUGGUGUUUAUUGC  
 UGAUUUAUUUAUAAUUGCCAGAUAUUUAUGGGUUGUUGCUUUGCUUGGAAUACUAGGAACAUUGAUGCU  
 ACUUCAACUGGUAAUUAUUUAUUUAUUUAUAGGUUAUCUUUAGACAUGGCAAGCUUAGGCCUUUUGAGAGACA  
 15 UAUCUAAUGUGCCUUUCUCCCCUGAUGGCAAACCUUGCACCCCACCUGCUCUAAUUGUUUAUUGGCCAUUAAA  
 UGAUUUAGGUUUUUUACACCACUACUGGCAUUGGCUACCAACCUUACAGAGUUGUAGUACUUUCUUUUGAACUU  
 UUUAAUGCACCGGCCACGGUUUGUGGACCAAAAUUAUCCACUGACCUUAAUUAAGAACAGUGUGUCAUUUUUA  
 AUUUUAAUGGACUCACUGGUACUGGUGUUUAAUCUUUCUCAAAGAGAUUUCAACCAUUUUAACAAUUUGG  
 CCGUGAUGUUUCUGAUUUUACUGAUUCCGUUCGAGAUCCUAAAACAUCUGAAAUAUUAGACAUUUACCUUGC  
 20 UCUUUUGGGGGUGUAAGUGUAAUUACACCUGGAACAAUUGCUUACUGAAGUUGCUGUUCUUAUUAAGAUG  
 UUAACUGCACUGAUGUUUCUACAGCAUUAUGCAGAUCAACUCACACCAGCUUGGCGCAUUAUUAUUCUACUGG  
 AAACAAUGUAUUCAGACUCAAGCAGGCGUGUCUUAUAGGAGCUGAGCAUGUCGACACUUCUUAUGAGUGCGAC  
 AUUCCUAUUGGAGCUGGCAUUGUGCUAGUUACCAUACAGUUUCUUUAUUACGUAGUACUAGCCAAAAUCUA  
 UUGUGGCUUAUACUAUGUCUUUAGGUGCUGAUAGUUCUAAUUGCUUACUCUAAUUAACACCAUUGCUUAUACCUAC  
 25 UAACUUUUAUUAUAGCAUUAUACAGAAUUAUGCCUGUUGUUUCUUAUGGCUAAAACCCUGAUAUUGUAAUUAUG  
 UACAUCUGCGGAGAUCUACUGAAUGUGCUAAUUGCUUCUCCAAUUAUGGUAGCUUUUGCACACAACUAAAUC  
 GUGCACUCUCGUGGUUUUGCUGUGAAGGAGUUGCAACACAC

## Translation 1

30 Nucleotides 3 to 818: Frame 3 272 aa (orf 1ab)  
 LQNYGENAVIPQGIMMNVAKYTQLCQYLNTLTAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLPTGTLVDSLDLNDVSDA  
 DSTLIGDCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKITEHSWNADLYKLMGHFS  
 WWTAFVTNVNASSSEAFLLIGANYLKPKEQIDGYTMHANYIFWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKENQIND  
 35 MIYSLLEKGRLLIIRENNRVVSSDILVNN

## Translation 2

40 Nucleotide 828 to 3089: Frame 3 756 aa (S protein)  
 MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPVI  
 PFKDGIYFAATEKSNVVRGWFVFGSTMNKSQSVIIINNSTNVVIRACNFELCDNPFPAVSKPMGTQHTMTIFDNFNCCTFE  
 YFSDAFSLDVSEKSGNEFKHLREFVFKNKDGLYVYKGYQPIDVVRDLPSGFNTLKFIFKPLGINITNFRAILTAFSPAQD  
 45 IWGTSAAAYFVGYLKPFTFMLKYDENGITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPBGDVRFPNITNLCP  
 FGEVFNATKEFSPVYAWERKKISNCVADYSVLNSTFFSTFKCYGVSAATKLNDLCFSNVYADSFVVGDDVRQIAPGQTGVI  
 ADYNYKLPPDFMGCVLAWNTRNIDATSTGNYNYKYRILRHGKLRPFERDISNVFPSPDGKPCPTPALNCYWLNDYGYFTT  
 TGIGYQPYRVVLSFELLNAPATVCGPKLSTDLIKNCVNFNFNLGTGTGLVTPSSKRFQFFQFGRDVSDFDTSVRDPKT  
 SEILDISPSCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTOAGCLIGAEHVDTSYEC  
 50 DIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMKTSVDCNMYICGDST  
 ECANLLQYGSFCTQLNRALSWYCC

## Alignment 1 of translation 1

55 replicase [bovine coronavirus]  
 Identities = 183/271 (67%),

Query: 3 LQNYGENAVIPQGIMMNVAKYTQLCQYLNTLTAVPYNMRVIHFGAGSDKGVAPGTAVLR 182  
 L NYG+ +P G MMNVAKYTQLCQYLNT TLAVP NMRV+H GAGS+KGVAPG+AVLR  
 60 Sbjct: 6822 LWNYGKPVTLPTGCMNVAKYTQLCQYLNTTTLAVPVNMRVLHLGAGSEKGVAPGSVAVLR 6881  
 Query: 183 QWLPTGTLVDSLDLNDVSDADSTLIGDCATVHTANKWDLIISDMYDPRTKHVTKENDSK 362  
 QWLP GT+LVD+DL FVSD+ +T GDC T+ +WDLIISDMYDPTK++ + N SK  
 65 Sbjct: 6882 QWLPAGTILVDNDLYPFVSDSVATYFGDCITLPFDCQWDLIISDMYDPTKNGEYNVSK 6941  
 Query: 363 EGFFTYLCGFIKQKLALGGSIAVKITEHSWNADLYKLMGHFSWWTAFVTNVNASSSEAF 542  
 +GFFTY+C I+ KLALGGS+A+KITE SWNA+LYKLMG+F++WT F TN NASSSE FL

## Fig. 2. Cont.

Sbjct: 6942 DGFFTYICHMIRDKLALGGSVAIKITEFSWNAELYKLMGYFAFWTVFCTNANASSSEGFL 7001

Query: 543 IGANYLGKPKKEQIDGYTMHANYIFWRNTNPIQLSSYSLEFDMSKFPLKLRGTAVMSLKENQ 722  
 IG NYLGKPK +IDG MHANY+FWRN+ +YSLEDM+KFPLKL GTAV++L+ +Q

5 Sbjct: 7002 IGINYLGKPKVEIDGNVMHANYLFWRNSTVWNGGAYSLEFDMAKFPLKLAGTAVINLRADQ 7061

Query: 723 INDMIYSLLEKGRLLIIRENNRVVSSDILVN 815  
 INDM+YSLEKGL++R+ N+ V D LVN

10 Sbjct: 7062 INDMVYSLEKGLLVRDNTKEVFVGDLSLVN 7092

Alignment 2 (Spike protein of coronavirus)  
 E2 glycoprotein precursor - murine hepatitis virus (strain JHM); contains  
 spike glycoprotein

15 Identities = 199/798 (24%), Positives = 314/798 (39%), Gaps = 48/798 (6%)  
 Frame = +3

Query: 828 MFIFLLFLTITSGSDLDRCITFDDVQAPNYTQHTSSM-----RGVYYP-DEI 965  
 +F+F+L L G D E +Q NY + +S RG YY D +

20 Sbjct: 2 LFVFILLPLSCLGYIGD----FRCIQTVNYNGNNASAPSISTEAVDVSKGRGTYVLDLV 57

Query: 966 FRSDTLYLTQDLFLPF----YSNV--TGFHTINHTFGNP--VIPFKDGIYFAATE-KSNV 1118  
 + + TL LT + P Y N+ TG +T++ T+ P + F DGI+ K+N

25 Sbjct: 58 YLNATLLLTG--YYPVDGNSYRNALALGTNTLSLTWFKPPLSEFNDGIFAKVQNLKTNT 115

Query: 1119 VRGW-----VFGSTMNNKXXXXXXXXXXXXXACNFELCDNPFPAVSKPMGTQHT 1277  
 G V GS N C + +C P+ KP

30 Sbjct: 116 PTGATSYFPTIVIGSLFGNTSYTVVLEPYNNIIMASVCTYTICQLPY-TPCKP----- 167

Query: 1278 MIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGLFYVY---KGYQPIDVVR 1448  
 N + + DV K R F F +LY + +G

35 Sbjct: 168 -----NTNGNRVIGFWHTDVKPPICLLK--RNFTFNVNAPWLYFHFYQQGGTFYAYYA 218

Query: 1449 DLPSEFNTLKPFLKPLGINITNFRAILTAFSPAQDIWGTSAAYFVGYLKPTTFMLKYD 1628  
 D PS L F + +G +T+ + +P T A Y+V L ++ ++

40 Sbjct: 219 DKPSATTFL---FSVYIGDILTQYFVLPFICTPTAG--STLAPLYWVTPLLKRQYLFNFN 273

Query: 1629 ENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSPGDVVR-FPNITNLCFPFG 1805  
 E G IT AVDC+ + ++E+KC +S G+Y S + V P G V R PN+ + C

45 Sbjct: 274 EKGVITSAVDCASSYISEIKCKTQSLPSTGVYDLSGYTVQPVGVVYRRVFNLPD-CKIE 332

Query: 1806 EVFNATKFPVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNLCFSNVYADS 1985  
 E A PS WER+ NC + S L + C + A+K+ +CF +V D

50 Sbjct: 333 EWLAKSVSPINWERRTFQNCNFNLSLLRYVQAESLSCNNIDASKVYGMCFGSVSVDK 392

Query: 1986 FVVKGDDVRQIAPGQTGVADIYNYKLPDDFMGCVLAWNTRNIDATSTGNYNKYRYLRHG 2165  
 F + + G +G + NYK+ C L ++ + T NYN R+G

55 Sbjct: 393 FAIPRSRQIDLQIGNSGFLQTANYKIDTAATSCQLYYSLEPKNNVT-INNYNPSSWNRRYG 451

Query: 2166 KLRPFERDISNVFSPDGKPCPTPALNCYWPLNDYGFYTTTGIGYQPYRVVLSFELLNA 2345  
 + +ND R + + LLN

60 Sbjct: 452 -----FKVND-----RCQIFANILLNG 468

Query: 2346 --PATVCGPKL---STDLIKNQCVNFENGLTGTGVLTTP-SSKRFPFQQFGRDVSDFTD 2507  
 T C L +T++ CV ++ G+TG GV + + +Q DV+ +

65 Sbjct: 469 INSGTTCSTDLQLPNTAVATGVCVRYDLYGITGQGVFKEVKADYNSWQALLYDVNGNLN 528

Query: 2508 SVRDPKTSEILDSPCSFGGVSIVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWR 2687  
 RD T++ I C G VS + E A+LY+++NC+ V T + + P

70 Sbjct: 529 GFRDLTTNKTYTIRSCYSGRVSAAY--HKEAPEPALLYRNINCSYVFTNNISREENPL-- 584

Query: 2688 IYSTGNNVFQTAGCLIGAEH--VDTSYECDIPIGAGICASYHTVSLLR---STSQK--S 2846  
 N F + GC++ A++ + C++ +GAG+C Y R ST + +

75 Sbjct: 585 -----NYFDSYLGCVVNADNRTDEALPNCNLRMGAGLCVDYSKRRARRSVSTGYRLTT 638

Query: 2847 IVAYTMSLGADSSIAYSN-NTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECA 3023  
 Y L DS + + IPTNF+I E + + K ++DC ++CGD+ C

80 Sbjct: 639 FEPYMPMLVNDVSVQSVGGLEYMQIPTNFTIGHHEEFIQIRAPKVITDCAAFVCGDNAACR 698

Query: 3024 NLLLQYGSFCTQLNRALS 3077

Fig. 2. Cont.

L++YGSFC +N L+  
 Sbjct: 699 QQLVEYGSFCDNVNAILN 716

**RDG1 seq**

5 UUCAAGGCcUCAAACNUAUGUAAACACAACAACUAAUCAGGGMUGcUGAAAUCHGSGCUUCUGCUAAUCUUGC  
 UGCUACUAAAAUGUCUGAGUGUGUUCUUGGACAAUAAAAAGAGUUGACUUUUGUGGAAGGGCUACCACCUU  
 AUGUCCUCCCCACAAGCAGCCCCGCAUGGUGUUGUCUCCUACAUGUCACGUAUGUGCCAUCCCAGGAGAGGA  
 10 ACUUCACCACAGCGCCAGCAAUUUGUCAUGAAGGCAAAGCAUACUCCCCUGUGAAGGUGUUUUUGUGUUUAA  
 UGGCACUUCUUGGUUUUAUUACACAGAGGAACUUCUUUUCUCCACAAUAAUACUACAGACAAUACAUUUGUC  
 UCAGGAAAUUGUGAUGUCGUUAUUGGCAUCAUUAACAACACAGUUUAUGAUCCUCUGCAACCUGAGCUUGACU  
 CAUUCAAAGAAGAGCUGGACAAGUACUCAAUUAUACAUCACCAGAUGUUGAUCUUGGCGACAUUUCAGG  
 CAUUAACGCUUCUGUCAACAUAUCAAAGAAUUGACCGCCUCAAUGAGGUCGCUAAAAUUUUAAUUGAA  
 UCACUCAUUGACCUUCAAGAAUUGGGAAAAUUGAGCAAUAUAUUAAGUGgCCCUGGUACGUCUGGGU

15 *Translation 1*  
 Nucleotides 3 to 650: Frame 3; 216 aa  
 QSLQXYVTQQLIRXAEIXSANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAIC  
 HEGKAYFPREGVVFVNGTSWFITQRNFFSPQIITDNTFVSGNCDVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPD  
 20 VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVW

*Translation 2*  
 Nucleotides 37 to 339: Frame 1; 101 aa  
 SGXLKXLLILLILLKCLSVFLDNQKELTFVERATTLCPSHKQPRMVLSSYMSRMCHPRRGTSPPQRQQFVMKAKHTSLVKV  
 25 FLCLMALLGLHRTSFLHK

*Translation 3*  
 Nucleotides 343 to 576: Frame 1; 78 aa  
 LLQTIHLSQEIVMSLLASLTTFMILCNLSLTHSKKSWTSTSKIHHQMLILATFQALTLLSSTFKKKLTASMRSLKI  
 30

*Alignment of translation 1*

S glycoprotein [murine hepatitis virus]  
 Length = 1376

35 Identities = 86/218 (39%), Positives = 129/218 (59%), Gaps = 3/218 (1%)  
 Frame = +3

Query: 6 SLQTYVTQQLIRXAEIXSANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVF 185  
 +L Y+++QL I SA A K++ECV Q+ R++FCG G H++S Q AP+G+ F  
 40 Sbjct: 1105 ALNAYISKQLSDSTLIKFSAAQAEKVNCEVKSQTRINFCGNGNHILSLVQNPAYGLYF 1164

Query: 186 LHVTVVPSQERNFTTAPAICHEG-KAYFPREGVVFVNGTSWFITQRNFFSPQIITDNTF 362  
 +H +YVP+ +P +C G + P+ G FV + W T +++ P+ IT N+  
 45 Sbjct: 1165 IHFSYVPTSFITANVSPGLCISGDRGLAPKAGYFVQDDGEWKFTGSSYYYPEPITDKNSV 1224

Query: 363 VSGNCDVIGIINNTVYDPLQPELDSFKEELDKYFKNHTS--PDVDLGDISGINASVVNI 536  
 V +C V + + P L FKEELDK+FKN TS PD+ L D +N + +++  
 Sbjct: 1225 VMSSCSVNYTKAPEVLLNSSIPNLPDFKEELDKWFKNQTSIAPDLSL-DFEKLNVTFDL 1283

50 Query: 537 QKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVW 650  
 E++R+ E K LNES I+L+E+G YE Y+KWPWYVW  
 Sbjct: 1284 SDEMNRQEAIKKLNESEYINLKEVGTIEMVYKWPWYVW 1321

**EMC-8**

55 AGGCCAAAACAGCGCCGACCCCAAGGUUUACCCAAUAAUACUGCGUCUUGGUUCACAGCUCUCACUCAGCAUG  
 GCAAGGAGGAACUUGAUUCCUCGAGGCCAGGGCGUCCAUAACACCAAUAGUGGUCCAGAUGACCAAU  
 UGGCUACUACCGAAGAGCUACCCGACGAGUUCGUGGUGGACGGCAAAUAGAAAGAGCUCAGCCCCAGAUGG  
 60 UACUUCUAUUAUACUAGGAACUGGCCAGAGCUACUUCUCCUACGGCGCUAACAAGAAGGCAUCGUUAGGG  
 UUGCAACUGAGGGAGCCUUGAAUACACCCAAAGACCACAUUGGCACCCGCAAUCCUAUAACAUAUGUUGCC

*Translation*  
 Nucleotides 1 to 363: Frame 1; 121 aa  
 RPKQRRPQGLPNNTASWFTALTQHGKEELRFPRGQGVPIINTNSGPDDQIGYYRRATRRVRGGDGKMKELSPRWYFYLLGTG  
 65 PEASLPYGANKEGIVVATEGALNTPKDHIGTRNPNNNKA

## Fig. 2. Cont.

## Alignment

nucleocapsid protein - bovine coronavirus (strain Mebus)

5     Identities = 55/129 (42%),

Query: 1    RPKQRRPQGLPNNTA-----SWFTALTQHGK-EELRFPRGQGVPIINTNSGPDQIGYYRR 162  
           +PKQ    LP+            SWF+ +TQ K +E F GQGVPI            + GY+ R

10    Sbjct: 44 QPKQTATSQLPSGGNVVPYYSWFSGITQFQKGKEFEAEGQGVPIAPGVPA TEAKGYWYR 103

Query: 163 ATRR-VRGGDGKMKELSPRWYFYLLGTGPEASLPYGANKEGIVWVATEGA-LNTPKDHIG 336  
           RR + DG ++L PRWYFYLLGTGP A   YG + +G+ WVA+ A +NTP D I

15    Sbjct: 104 HNRRSFKTADGNQRQLLPRWYFYLLGTGPHAKDQYGTIDGVFWVASNQADVNTPAD-IL 162

Query: 337 TRNPNNNXA 363  
           R+P+++ A

15    Sbjct: 163 DRDPSSDEA 171

EMC-11: unknown sequence

20    UUGCAUACCGCAAUGUUCUUCUUCGUAAGAACGGUaAUAAGGGAGCCGGUGGUCAUAGCUgUGGCAUGAUCUA  
       AAGUCUUAUGACUUAAGGUGACGAGCUUGGCACUGAUCCCAUUGAAGAUUAUGAACAAAACUGGAACACUAAGC  
       AUGGCAGUGGUGCACUCCGUGAACUCACUCGUGAGCUCUAAUGGAGGUGCAGUCACUCGCUAUGUCGACAAACAA  
       UUUCUGUGGCCCCAGAUAGGUAACCCUCUUGAUUGCAUCAAAGAUUUUCUCGCACGCGCGGGCAAGUCAUUGUC  
       ACUCUUUCCGAACAACUUGAUUACAUCGAGUCGaAGAGAGGUGUCUACUGCUGCCGUGACCAUGAGCAUGAAA

25    UUGCCUgGGUUCACUGAGCGCUCUGAUAAAGAGCUACGAGCACCAGACACCCUUCGaAAUUAAGAGUGCCAAGA  
       AAaUUGACACUUUCAAAGGGGAUGCCCCAAAGCUUGUGUUUCCUCUUAACUCAAAGUCAAGUCAUUCAA  
       CCACGUGUUGAAAAAGAAAAAGACUGAGGGUUAUACUGGGGCUAUACGCUCUGUGUACCCUGUUGCAUCCAC  
       AGGAGUGUAACAAUAGCACUUGUCUACCUUGAUGAAUGUAUUAUGCGAUGAAGCUUAUGGCAGACGUG  
       CGACUUUCUGAAAGCCACUUGUGAACAUUGGGCACUGAAAAUUAGUUAUUGAAGGACCUAGUACAUGUGGG

30    UACCUACCUACUAAUGCUGUAGUGAAAAUGCCAUGUCCUGCCUGUCAAGACCCAGAGAUUGGACCUGAGCAUA  
       GUGUUGCAGAUUAUCACAACCACUCAACAUUGAAACUCGACUCCGCAAGGGAGGUAGGACUAGAUGUUUUGG  
       AGGUGUGUGUUUGCUAUGUUGGCUGCUAUAUAAGCGUGCCUACUGGGUUCUUGGCUAGUGCUGAUAAU  
       GGCUCAGGCCAUACUGGCAUUAUGGUGACAAUGGAGACCUUGAAUGAGGAUCCUUGAGAUACUGAGUC  
       GUGAACGUGUUAACAUUAACAUUGUUGGCGAUUUUCAUUUGAAUGAAGAGGUUGCCAUCAYUUUGGCAUCYUU

35    CUCUGCUUCUACAAGUCCUUUAUUGACACUAAUAAAGAGUCUUGAUUACAAGUCUUUCAAACCAUUGUUGAG  
       UCCUGCGGUAACUAAUAAAGUUACCAAGGGAAAGCCCGUAAAGGUGCUUGGAACAUUGGACAACAGAGUCAG  
       UUUUAACACCACUGUGUGGUUUUCCUCACAGGCUGCUGGUGUUAUCAGAUCAAUUUUUGCGCGCACACUUGA  
       UGCAGCAAACCACUCAAUUCUGAAUUGCAAAGAGCAGCUGUACCAUACUUGAGUUAUUUCUGAACAGUCA  
       UUACGUCUUGUCGACGCCAUGGUUUUAUACUUCAGACCUGCUCACCAACAGUGUCAUUUAUUGGCAUAUGUAA

40    CUGGUGGUCUUGUACAACAGACU

## Translation of putative open reading frames

45    >~out: 78 to 1: Frame -2           26 aa  
       DFRSCHSYDHLRPYYRSYEEHCGMQ

      >~out: 59 to 379: Frame 2           107 aa  
       LWHDLSYDLGDELGTDPIDYEQNWNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLS  
       EQLDYIESKRGVYCCRDHEHEIAVH

50    >~out: 283 to 89: Frame -1           65 aa  
       LARACEKIFDAIKRVPWIWATEIIVDIASDCTSIELTSEFTECTAMLSPVFLFIIFNGISAKLVT

      >~out: 90 to 614: Frame 3           175 aa  
       VTSLALIPKIMNKTGTLMAVHVSNSLVSSMEVQSLAMSTTISVAQMGTLLIASKIFSHARASQCALFPNNLITSSRRE  
       VSTAATMSMKLPFGFTERSDKSYEHQTPFEIKSAKKIDTFKRGMPQSLCFLLTQKSFSFNHVLKRRLRVSWGVYALCTLL  
       HLHRSVTICTCLP

55    >~out: 204 to 124: Frame -2           27 aa  
       RVTAPPLSSRVSSRSAPLPCLVFQFCS

      >~out: 312 to 208: Frame -2           35 aa  
       SSCSERVHIDLPARARKSLMQSRGYPSGPQKLLST

60    >~out: 485 to 258: Frame -3           76 aa  
       EETQALGHSPFESVNFGLTGNFEGCLVLVALIRALSEPRQFAHGHGSSRHLSSSTRCNQVVRKECTLTCPVRRENL

      >~out: 397 to 287: Frame -1           37 aa  
       LLSERSVNPNGFMMLMVTAAVDTSRLDLVIKLFGKSAH

      >~out: 364 to 486: Frame 1           41 aa  
       NCLGSLALIRATSTRHPSKLRVPRKLTLSKGECPKACVSS

65    >~out: 490 to 401: Frame -1           30 aa  
       VKRKHLWGIPLLKVSIFLALLISKGVWCS

      >~out: 446 to 1483: Frame 2           346 aa

Fig. 2. Cont.

HFQKGNAPKLVFPLNSKVVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEASWQTCDFLKATCEHC  
 GTENLVIEGPSTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSNIETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVP  
 RASADIGSGHTGITGDNVETLNEDLLEILSRERVNINIVGDFHLNNEEVAIXLAXFSASTSAFIDTIKSLDYKSEFTIVESC  
 5 GNYKVTKGKPKVGAWNIGQQRSVLTPICGFPSQAAGVIRSIFARTLDAANHSIPDLQRAAVTILDGISEQSLRLVDAMVYT  
 SLLTNSVIIMAYVTGGLVQQT  
 >~out: 643 to 494: Frame -1 50 aa  
 SFIAMITFHQGRQVHIVTLLWRCNRVHRAITPHETLSLFLFNTWLNDDFD  
 >~out: 627 to 511: Frame -2 39 aa  
 LHFIVKDKCILLHSCGDATGYTERIRPMKPSVFFSTRG  
 10 >~out: 704 to 612: Frame -3 31 aa  
 LNFQCHNVHKWLSESRTSAMKLHRNDYISSR  
 >~out: 774 to 631: Frame -2 48 aa  
 QAGHGIFTALVGRYPHVLGPSITKFSVPQCSQVAFRKSHVCHASSQ  
 >~out: 826 to 737: Frame -1 30 aa  
 15 VVVIICNTMLRSNLWVLTGRTWHFHYSSR  
 >~out: 863 to 744: Frame -3 40 aa  
 SYLPCGVFEQCLSGCDNLQHYAQVQSLGLDRQDMAFSLQH  
 >~out: 756 to 992: Frame 3 79 aa  
 KCHVLPVKTRQLDLSIVLQIITTTQTLKLDARSVEGLDVLEAVCLPMLAAIISVPTGFLVLVLILAQAILALLVTMWRP  
 20 >~out: 952 to 830: Frame -1 41 aa  
 ANISTSTRNPVGTLLIAANIGKHTASKTSSPTSLSAESSFNV  
 >~out: 1056 to 922: Frame -2 45 aa  
 KSPTMLMLTRSLRSRSSFVKVSTLSPVMPVWPEPISALARGTQ  
 >~out: 1237 to 956: Frame -1 94 aa  
 25 SLLSNVPSTFYGLSLGNFIVTAGLNNGFERLVIKTLYSVNKGTCRSREXCQXDGNLFIQMKIANNVNVNTFTTQYLKEILI  
 QGLHIVTSNAGMA  
 >~out: 1140 to 1060: Frame -2 27 aa  
 SRLFIVSIKALVEAEXDAKXMATSSFK  
 >~out: 1131 to 1205: Frame 3 25 aa  
 30 RVLITSLSKPLLSPAVTIKLPRESP  
 >~out: 1410 to 1183: Frame -2 76 aa  
 TMASTRRNDCEIPSSMVTALCKSGIEWFAASSVRKIDLITPAACEGKPHSGVKTDLCCPMFQAPFTGFPLVTL  
 >~out: 1186 to 1311: Frame 1 42 aa  
 SYQGKARKRCLEHWTEISFNTTVWFSLTGCWCYQINFCAHT  
 35 >~out: 1283 to 1191: Frame -3 31 aa  
 HQQPVRENHTVVLKLISVVQCSKHLLRAFPW  
 >~out: 1248 to 1457: Frame 3 70 aa  
 HHCVVFPRLVLSDQFLRAHLMQQTQFLICKEQLSPYLMVFLNSHYVLSTPWFILOTCSPTVSLLWHM  
 >~out: 1381 to 1482: Frame 1 34 aa  
 40 TVITSCRRHGLYFRPAHQQCHYYGICNWWSCSTD

**EMC12:** unknown sequence

UGCUUGCUCAUGCUGAAGAGACAAGAAAAUUAUUGCCUAUAUGCAUGGAUGUAGAGCCAU  
 AAUGGCAACCAUCCAACGUAAGUAUAAAGGAAUUAUAAUUAAGAGGGCAUCGUUGACUUAU  
 45 GGUGUCCGAUUCUUCUUAUUAUACUAGUAAAGAGCCUGUAGCUUCUUAUUAUACGAAGCUGA  
 ACUCUCUAAAUGAGCCGCUUGUCACAAUGCCAAUUGGUUAUGUGACACAUGGUUUUAAUCU  
 UGAAGAGGCUGCGCGCUGUAUGCGUUCUCUUAAGCUCUGCCGUAGUGUCAGUAUCAUCA  
 CCAGAUGCUGUUAUACAUAAUAAUGGAUACCUCACUUCGUCAUCAAAGACAUCUGAGGAGC  
 ACUUUGUAGAAACAGUUUCUUUGGCUGGCUCUUAACAGAGAUUGGUCCUAUUCAGGACAGCG  
 50 UACAGAGUUAGGUGUUGAA

*Translation of putative open reading frames*

>~out: 3 to 446: Frame 3 148 aa  
 LAHAETRKLMPICMDVRAIMATIQRKYKGIKIQEGIVDYGVREFFYTSKEPVASIIITKLNSLNEPLVTMPIGYVTHGFNL  
 55 EEAARCMRSLKAPAVSVSSPDVATTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSQORTELGVE  
 >~out: 100 to 11: Frame -2 30 aa  
 ILIPLYLRWMVAIMALTSMHIGINFLVSSA  
 >~out: 188 to 33: Frame -1 52 aa  
 RVQLRNNRSYRLFTSIKESDTIVNDALLNFNSFILTLDGCHYGSNIHAYRH  
 60 >~out: 64 to 159: Frame 1 32 aa  
 WQPSNVSIELKFKRASLTMVSDSSFILVKSL  
 >~out: 220 to 143: Frame -2 26 aa  
 PIGIVTSGSFREFSFVIEATGSLV  
 >~out: 293 to 192: Frame -1 34 aa  
 65 HYGRSFKRTHRTARSLFKIKTMCHITNWHCDKRLI  
 >~out: 397 to 224: Frame -2 58 aa  
 EPAKETVSTKSSDVFDDEVRYPLYVVTASGDPTDTTAGALRERIQRASSRLKPCVT  
 >~out: 229 to 288: Frame 1 20 aa



Fig. 2. Cont.

```
HMVLILKRLRAVCVLLKLLP
>-out: 292 to 372: Frame 1      27 aa
CQYHHQMLLLHIMDTSLRHQRHLRSTL
>-out: 444 to 340: Frame -3    35 aa
5 QHLTLYAVLNRTNLCKSQPKKLFQLQSAPQMSLMTK
>-out: 416 to 351: Frame -1    22 aa
IGPISVRASQRNCFYKVLLRCL
>-out: 365 to 445: Frame 2     27 aa
10 GALCRNSFFGWLLQRLVLFRTAYRVRC
>-out: 376 to 435: Frame 1     20 aa
KQFLWLALTEIGPIQDSVQS
```

15

Figure 3.

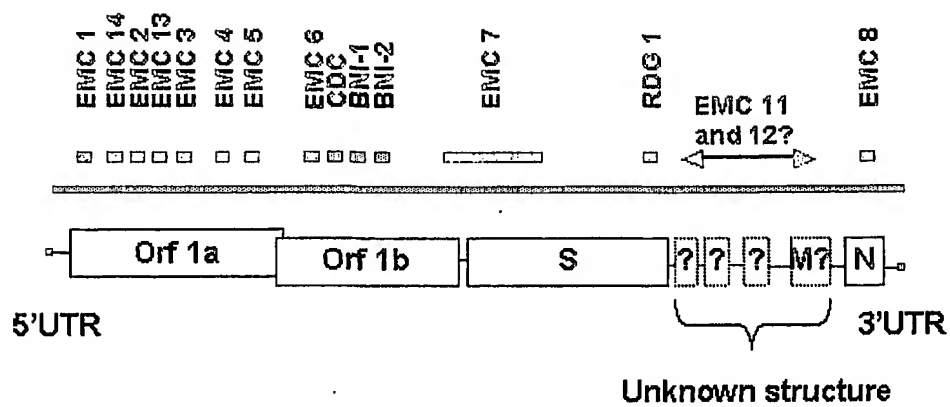


Figure 4.

Comparison of N-termini of the S proteins of the group 2 coronaviruses

5

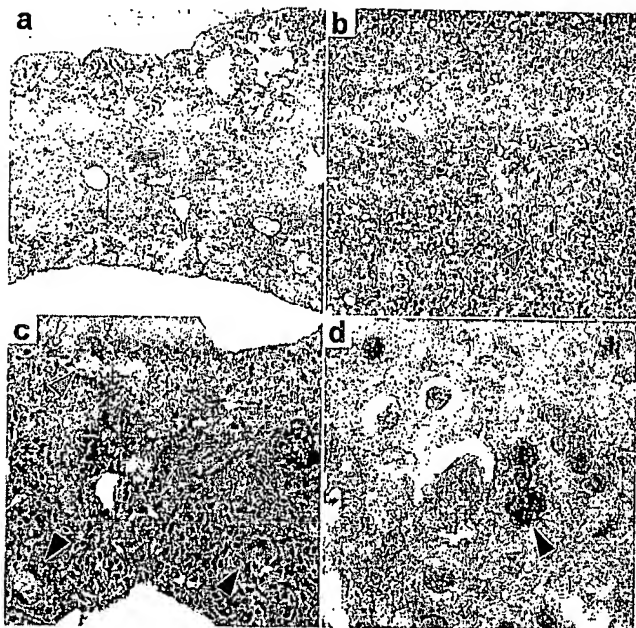
HCV OC43	MFLILLISLPTAFAVIG <b>DL</b> - <del>KCTTVS</del> INDID
MHV A59	MLFVFILEFLPSCCLGYIG <b>DF</b> - <del>RCIQLVNS</del> NGA
BCV	MFLILLISLPMAFAVIG <b>DL</b> - <del>KCTTVS</del> INDVD
SARS	MF-IFLLEFL-TLTSG-SDLDRCTTFDDVQAP

10

Figure 5.



Figure 6.



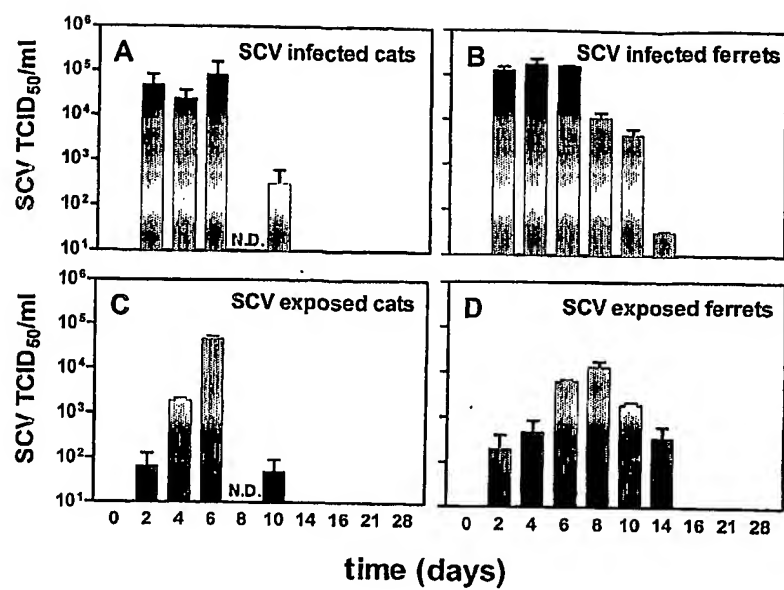


Figure 7.

Figure 8 Detection of SCV in postmortem tissues of experimentally SCV infected cats and ferrets

Tissue	Cats				Ferrets			
	1	2	3	4	1	2	3 <sup>†</sup>	4
Trachea	+/	+/	+/	+/	+/	+/	+/	+/
Lung	++*	+	+	+	+	+	+	+
Tracheo-bronchial lymph node	+/	+/	+/	+/	+/	+/	+/	+/
Duodenum	-/-	-	+	-	+	+	+	+
Jejunum	-/-	-/-	-	+/	-	+	-/-	-
Ileum	-/-	+	+	-/-	-	+	+/	-
Mesenteric lymph node	-/-	+	-/-	-/-	+	-	-/-	-
Kidney	-/-	+	+	-/-	+	-	-/-	-
Urinary bladder	-/-	+	-/-	-/-	-	-	-	-
<b>Peripheral blood mononuclear cells</b>	-/-	-	-	-/-	-	-	-	-

\* virus isolation/real time PCR results are depicted at day 4 p.i.

<sup>†</sup> died during the experiment.

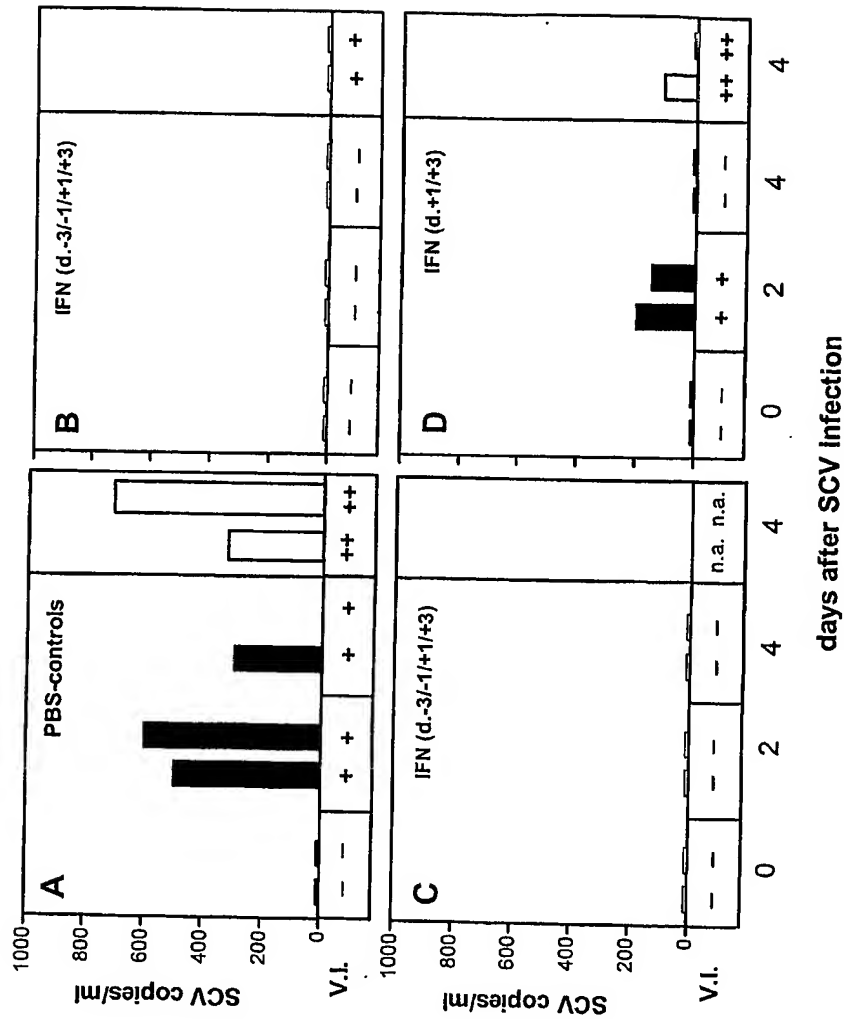


Figure 9



Figure 10

1 atattaggtt ttaccctacc caggaaaagc caaccaacct cgatctcttg tagatctgtt  
 61 ctctaaacga actttaaat ctgtgtagct gtgcctcgcc tgcctgccta gtgcacctac  
 121 gcagataaaa caataataaa ttttactgtc gttgacaaga aacagagtaac tgcctcctct  
 181 tctgcagact gcttacgggt tgcctcgtgt tgcagtcgat catcagcata cctaggtttc  
 241 gtccgggtgt gaccgaaagg taagatggag agccttgttc ttggtgtcaa cgagaaaaca  
 301 cagctccaac tcagtttgcc tgcctctcag gttagagacg tgcctagtcg tggcttcggg  
 361 gactctgtgt aagaggccct atcggaggca cgtgaacacc tcaaaaatgg cacttctggt  
 421 ctatagagc tggaaaaagg cgtactgcc cagcttgaac agcctatgt gttcattaaa  
 481 cgtctgatg ccttaagcac caatcacggc cacaaggctg ttgagctggt tgcagaaatg  
 541 gacggcattc agtacggctg tagcgggata acactgggag tactcgtgcc acatgtgggc  
 601 gaaaccccaa ttgcataccg caatgttctt cttcgtaaag acggtataaa gggagccggg  
 661 ggtctagct atggcatcga tctaaagtct tatgacttag gtgacgagct tggcaactgat  
 721 cccattagag attatgaaca aaactggaac actcgtatg tgcacaacaa tttctgtggc  
 781 ctcaactcgt agctcaatgg aggtgcagtc actcgtatg tgcacaacaa tttctgtggc  
 841 ccagatgggt accctcttga ttgcatcaaa gatitctctg cagcgccggg caagtcaatg  
 901 tgcactcttt ccgaacaact tgattacatc gatttctctg cagcgccggg caagtcaatg  
 961 gaccatgagc atgaaattgc ctggttctact gacgctctg ataaagacta cgagcaccag  
 1021 acaccctctg aaattaaagg tgccaaagaa ttgacacct tcaaaaggga atgcccacaa  
 1081 tttgtgttct ccttaactc aaaaagtcacaa gtcatccctg ttgcatctcc acaggaggtg  
 1141 actgaggggt tcatggggcg tatacgtctc ctgtacatc ttgcatctcc acaggaggtg  
 1201 aacaatatgc acttgtctac cttgatgaaa tgtaatcatt gcgatgaagt ttcattggcag  
 1261 acgtgggact ttctgaaagc cacttgtgaa cattgtggca ctgaaaattt agttattgaa  
 1321 ggactacta catgtgggta cctaccctact aatgctgtag tgaaaatgcc atgtcctgcc  
 1381 tgtcaagacc cagagattgg acctgagcat agtgttgcaag attatcacaa ccaactcaac  
 1441 attgaaactc gactccgcaa gggaggttagg actagatgtt ttggaggctg tgtgtttgcc  
 1501 tatgttggt gctataataa gcgtgcctac tgggttctc gtgctagtgc tgatattggc  
 1561 tcaggccata ctggcattac tggtagaat gtggagacct tgaatgagga tctccttgag  
 1621 taactgagtc gtgaacgtgt taacattaac attgttggcg atttctattt gaatgaagag  
 1681 gttggcatca ttttggcatc tttctctgct tctacaagt ccttatttga cactataaag  
 1741 agtcttgatt acaagtcttt caaaacatt gttgagtcct gcgtaacta taaagttaac  
 1801 agggaaagc ccgtaaaagg tgcttgaac attggacaac agagatcagt ttaaacacca  
 1861 ctgtgtgggt ttccctcaca ggtgtctgggt gttatcagat caatttttgc ggcacacct  
 1921 gatgcagcaa accactcaat tctgattttg caaagagcag ctgtcaccat acttgatggt  
 1981 attctgaac agtcattacg tctgtctgac gccatggtt atacttcaga cctgtctacc  
 2041 aacagtgtca ttattatggc tctgtatgaac ggtggtcttg tacaacagac tctcagtg  
 2101 ttgtctaac ttttgggcac tactgtttaa aaactcaggc ctatcttga atggattgag  
 2161 gcgaacatta gtgcaggagt tgaatttctc aaggatgctt gggagattct caaatttctc

Fig. 10. Cont

2221 attacagggtg tttttgacat cgtcaagggt caaatcacgg ttgtttcaga taacatcaag  
 2281 gattgtgtaa aatgcttcat tgatgttgtt aacaaggcac tcgaaatgtg cattgatcaa  
 2341 gtcaactatcg ctggcgcaaa gttgcgatca ctcaacttag gtgaagtctt catogctcaa  
 2401 agcaaggagac tttacacgtca gtgtatacgt ggcaaggagc agctgcaact actcatgcoct  
 2461 cttaaggcac caaagaagt aacctttctt gaagtgattt cacatgacac agtacttiacc  
 2521 tctgaggagg ttgtttctaa gaaacgtttaa ctgcagcac cgttgatagc  
 2581 ttcaaaaatg gagctatcgt tggcacacca ctctgtgtaa atggcctcat gctcttagag  
 2641 attaaaggaca aagaacaata ctgcgattg tctcctggtt tactggctac aaacaatgtc  
 2701 tttcgccttaa aagggggtgc accaatlaaa ggtgtaacct ttggagaaga tactgtttgg  
 2761 gaagtccaag gttacaagaa tftgagaatc acatttgagc ttgatgaacg tgttgacaaa  
 2821 gtgcttaatg aaaagtgttc tgtctacact gttgaatcgg gtaccgaagt tactgagttt  
 2881 gcatgtgttg tagcagaggc ttgtgtgaag actttacaac cagtttctga tctccttacc  
 2941 aacatgggtg ttgatcttga tgaftggagt gtactacat tctacttatt tgatgatgct  
 3001 ggtgaagaaa acttttcatc acgtatgtat tgttcttttt accctccaga tgaggagaag  
 3061 gaggacgatg cagagtgtga ggaagaagaa attgatgaaa cctgtgaaca tgaftacggt  
 3121 acagaggatg attatcaagg tctcctctcg gaatttggtg cctcagctga aacagttcga  
 3181 ttgagggaag aagaagagga agactggctg gatgatacta ctgagcaatc agagattgag  
 3241 ccagaaccag aacctacacc tgaagaacca gttaatcagt ttacttggtt tttaaaaactt  
 3301 actgacaatg ttgccattaa atgtgttgac atcgtttaagg aggcacaaag tgctaactct  
 3361 atggtgattg taaatgctgc taacatacac ctgaacacatg gtggtggtgt agcagggtgca  
 3421 ccacaacagg caaccaatgg tgccatgcaa aaggagagtg atgattacat taagctaaat  
 3481 gcccctctta cagttaggag gtcttgtttg cttctctggac ataacttgc taagaagtgt  
 3541 ctgcatgttg ttggacctaa cctaaatgca ggtgaggaca tccagcttct taaggcagca  
 3601 tatgaaaaat tcaattcaca ggaactctta ctgtaccat tgttgcagc aggcataatt  
 3661 ggtgctaacc cacttcagtc tttaacaagt tgcgtgcaga cggttcgtac acaggtttat  
 3721 attgcagta atgacaaagc tctttatgag caggttgtca tggattatct tgataacctg  
 3781 agccctagag tggaaagcacc taacaaagag gagccacca acaacaaga ttccaaaact  
 3841 gaggagaat ctgtcgtaca gaagcctgtc gatgtgaagc caaaatttaa ggcctgcatt  
 3901 gatgaggtta ccacaacact ggaagaaact aagtttctta ccaataagtt actcttgtt  
 3961 gctgatatca atggtaaagt ttaccatgat tctcagaaca tgcctagagg tgaagatatg  
 4021 tcttctcttg agaaggatgc acctacatg ttaggtgatg ttataactag tggtgatatac  
 4081 acttctgttg taataccctc caaaaaggct ggtggcacta accagttacc ctggacaagg atgtgtggt  
 4141 ttgaagaaag tggcagttga tgagtataa aagaaatgca aatctgcatt ttatgtacta  
 4201 tatacactg aggaagctaa gactgtctt taaggaaag atttctaggaa ctgtatcctg gaatttgaga  
 4261 ccttcagaag cactaatgc ctcatgctga agagacaaga aaattaatgc ctatatgcac ggaattgaga  
 4321 gaaatgttg caaccatcca acgttaagtat aaaggaatta aattcaaga gggcatcggt  
 4381 gactatgttg tccgattctt cttttactac agtaagagc ctgtagcttc tattattacg  
 4441 gactatgttg tccgattctt cttttactac agtaagagc ctgtagcttc tattattacg  
 4501 aagctgaact ctctaaatga gccgcttgc acaatgccaa ttggttatgt gacacatggt  
 4561 tttaatcttg aagaggctgc gcgctgtat cgttctctta aagctcctgc cgtagtgtca  
 4621 gtatcatcac cagatgctgt tactacatat aatggataacc tcaactcgtc atcaagaca

5

10

15

20

25

30

35

40

Fig. 10. Cont.

4681 tctgaggagc actttgtaga aacagtttct ttggtggct cttacagaga ttggtcctat  
 4741 tcaggacagc gtacagagtt aggtgttgaa tttcttaagc gtgtgacaa aattgtgtac  
 4801 cacactctgg agagccccgt cgagtttcat cttgacgggt aggtttttc acttgacaaa  
 4861 ctaaagagtc tcttatccct gcgggaggtt aagactataa agtgttcac aactgtggac  
 4921 aacactaatc tccacacaca gcttgtggat atgtctatga catatggaca gcagtttggg  
 4981 ccaacatact tggatgggtc tagtgttaca aaataaaac ctcagtataa tcatgagggg  
 5041 aagactttct ttgtactacc tagtgagttt tcttggtagg tacatgtctg cttaaagctt cgagtactac  
 5101 catactcttg atgagagttt ctcaagtttg tggtttaact tcaattaaat gggctgataa caattgttat  
 5161 tggaaaattc ttttattagc acttcaacag cctgctggt gatgctgcta actttgtgc actcatactc  
 5221 ttgtctagtg ttttattagc acttcaacag cctgctggt gatgctgcta actttgtgc actcatactc  
 5281 caagaggctt attatagagc cctgctggt tggcgagctt ggtgatgtca gagaaactat gacctatctt  
 5341 cttacagta ataaaactgt ttgctgagc atctgcaaag cgagttctta atgtgtgtg taaacattgt  
 5401 ctacagcatg ctaatttggg ctactacctt aacgggtgta gaagctgtga tgtatatggg tactctatct  
 5461 ggtcagaaaa ttaagacagg tttttccatt ccatgtgtgt ggtgtgtgta tgcatacaaa  
 5521 tatgataatc aacaagagtc ttcttttggg atgatgtctg caccacctgc tgagtataaa  
 5581 tatctagtagc gtacattctt atgtgcgaat gagtacactg gtaactatca gtgtgggtcat  
 5641 ttacagcaag taactgctaa ggagacctc agtactgat gttttctaca aggaacatc ttacataca  
 5701 taactgctaa taactgctaa ggagacctc agtactgat gttttctaca aggaacatc ttacataca  
 5761 atgtcagagt acaaggacc agtactgat gttttctaca aggaacatc ttacataca  
 5821 accatcaagc ctgtgtcgt taaactcgt tactatacag agcagcctat agacctgtga  
 5881 ttggatgggt attataaaa ggataatgct gataatttca aactcacatg ttctaacaca  
 5941 ccaactcaac cattaccaaa tgcgagttt tcaaatgaca ggtctacaa agccagcttc acgagagcta  
 6001 aaatttgcgt atgattttaa ctgaaatggc gatgtagtgg ctattgacta tagacactat  
 6061 tctgtcacat tcttcccaga tctgaaatggc gatgtagtgg ctattgacta tagacactat  
 6121 tcaggaggtt tcaagaaagg tgcataatga ctgcataagc caattgtttg gcacattaac  
 6181 cagggtacaa ccaagacaac gttcaaaaca aacacttggg gtttacgttg tctttggagt  
 6241 acagagccag tagatacttc aaattcattt gaagtctctg aagaagtagt ggaaaatcct  
 6301 atggacaatc ttgcttgtga aagtcaaca cccacctctg aagaagtagt ggaaaatcct  
 6361 accatacaga aggaagtcat agagtgtgac gtgaaaacta ccgaagtgtt aggcaatgtc  
 6421 taacttaaac catcagatga aggtgtttaa gtaacacaag agttaggtca tgaggatctt  
 6481 atggtgtcct atgtggaaaa cacaagcatt accattaaga aacctaatga gctttcacta  
 6541 gccttaggtt taaaacaaat tgcactcat ggtattgtcg caattaatag tgttcccttg  
 6601 agtaaaaatt tggcttatgt caaacattc ttaggacaag tgccttatgt gtttacta  
 6661 tgggctaaga gattagcaca acgtgtgttt acaattata gacttatagt gtttacta  
 6721 ttgttccaat tgtgtacttt tactaaaagt accaattcta gaattagagc ttactacctt  
 6781 acaactattg ctaaaaaatag tgttaagagt gttgctaaat tatgtttgga tgcgggacct  
 6841 aattatgtga gtcaccccaa attttctaaa ttgttcacaa tcgctatgtg gctattgttg  
 6901 ttaagtattt gcttaggttc tctaactctgt gtaactgtcg cttttgggtg actcttatct  
 6961 aattttgggt ctccttctta ttgtaatggc gttagagaat tgtatcttaa ttcgtctaac  
 7021 gttactacta tggatttctg tgaaggttct tttccttgca gcatttgggt aagtggatta  
 7081 gactcccttg attcttatcc agtcttggaa accattcaagg tgacgatttc atcgtaacag

Fig. 10. Cont.

7141 ctgacttga caatttttag tctggccgct gagtggggtt tggcatatat gttgttcaca  
 7201 aaattctttt atttattagg tctttcagct ataattcagg tgttttttgg ctattttgct  
 7261 agtcatttca tcagcaattc ttggctcagc gatgtacatc tctttttatca ttagtattgt acaaatggca  
 7321 cccgtttctg caatggtttag ggttgcaacc tcttcgactt gcatgatgtg ctataagcgc  
 7381 agctatgttc atatcatgga gtgtacaact atgtttaatg gcataagag atctttctat  
 7441 aatcgtgcca caccgcttga gtgtacaact atgtttaatg gcataagag atctttctat  
 7501 gtctatgcaa atggaggccg tggcttctgc aagatcaca attggaattg tctcaattgt  
 7561 gacacattt gcactggtag tacattcatt agtgatgaag ttgctcgtga ttgtcactc  
 7621 cagtttaaaa gaccaatcaa cctactgac cagtcatcgt atattgttga tagtgttgc  
 7681 gtgaaaaatg gcgccttca cctctacttt gacaaaggctg gtaaaaagac ctatgagaga  
 7741 catcgcctc cccattttgt caatttagac aatttgagag ctaacaacac taaaggltca  
 7801 ctgacctta atgtcatagt ttttgatgac aatttgagag ctaacaacac taaaggltca  
 7861 tctgctctg gtactacag tcagctgatg tgccaacctt tctgttgcct tgaccttaag  
 7921 cttgtatcag acgttggaga tagtactgaa gtttccgtta agatgtttga tgcctatgtc  
 7981 gacaccttt cagcaacttt tagtgttcc ttagatggtg tcccttctac attcgtgtca  
 8041 gctcacagcg agttagcaaa ggtgtagct ttagatggtg tcccttctac attcgtgtca  
 8101 gctgccgac aggtgttgc ttagatggtg ttagatggtg tcccttctac attcgtgtca  
 8161 aaactttcac atcacttga cttagaagt acaggtgaca gttgtaacaa tttcatgtctc  
 8221 acctataata aggttgaata catgacgccc agagatcttg gcgcatgtat tgcactgtaac  
 8281 gcaaggcata tcaatgcccc agtagcaaaa agtcacaatg ttctactcat ctggaatgta  
 8341 aaagactaca tgtctttatc tgaacagctg cgtaaacaaa ttcgtagtgc tgccaagaag  
 8401 acaacatac cttttagact aacttgtgct acaactagac aggttgtcaa tgcataaact  
 8461 actaaatct cactcaaggg ttgtaagatt gttagtactt gttttaaact tatgtttaag  
 8521 gccacattat tgtgcgttct tgctgcattg gtttgttata tggttatgcc agtacataga  
 8581 ttgtcaatcc atgatggtta cacaatgaa atcattggtt acaaaagccat tcaggatggt  
 8641 gtcactcgtg acatcatttc tactgatgat tgttttgcaa ataaacatgc tggttttgac  
 8701 gcatcgttga gccagcgtgg ttgttcatac aaaaatgaca aaagctgccc tgtagttagct  
 8761 gctatcatta caagagagat ttgtttcata gtgcctggct taccgggtac tgtgctgaga  
 8821 gcaatcaatg gtgacttctt gcattttcta cctcgtggtt ttagtgctgt tggcaacatt  
 8881 tgtacacac ctccaaact cattgagat agtggaact ctacctctgc tggcgttctt  
 8941 gctgctgagt gtacaaattt taaggatgct atgggcaaac ctgtgccata ttgttatgac  
 9001 actaatgtgc tagagggttc tatttcttat agtgagcttc gtccagacac togttatgtg  
 9061 cttatggatg gttccatcat acagtcttcc tagatggtga catgagcttc tggagggttc tgttagagta  
 9121 gtaacaaatt ttgatgctga gtaggggttg tagatgggtt cttataatg gcgaaaggctc agaagtaggt  
 9181 atttgcctat ctaccagtgg tagatgggtt tagatgggtt cttataatg gcgaaaggctc agaagtaggt  
 9241 ggaattttct gtggtgttga tgcgatgaat ctcatagcta acatctttac tctcttctg  
 9301 caactctggt gtgcttttaga ttgtgtcgtc tcaagatggt ctggtggtat tattgccata  
 9361 ttggtgactt gtgtgccta ctactttatg aaattcagac gtgttttttg tgagtacaac  
 9421 catgttgttg ctgctaagtc acttttgtt ttgatgtctt tcaactatct ctgtctggtg  
 9481 ccagcttaca gctttctgcc gggagtctac tcagtcttct acttgactt gacattctat  
 9541 ttcaccaatg atgtttcatt ctgggctcac ctccaatggt ttgccatggt ttctcctatt

Fig. 10. Cont.

9601 gtgccttttt ggataacagc aatctatgta ttctgtatct ctctgaagca ctgccattgg  
 9661 ttcttttaaca actatcttag gaaaagagtc atgtttaatg gagttacatt tagtaccttc  
 9721 gaggaggtcg cttttgtgac ctttttgcgc aacaaggaaa tgtacctaaa attgogtagc  
 9781 gagacactgt tgcacactac acagtataac aggtatcttg cttctatata caagtacaag  
 9841 tatttcagtg gacccctaga tactaccagg tatctgtaag cagcttgctg ccacttagca  
 9901 aaggctctaa atgacttttag caactcaggt gctgatgttc ttaccaaac ccacagaca  
 9961 tcaatcactt ctgctgttct gcagagtggt tttaggaaaa tggcattccc gtcaggcaaa  
 10021 gttgaagggt gcatgggtaca agtaacctgt ggaactacaa ctttaattgg attfggttg  
 10081 gatgacacag tatactgtcc aagacatgac atttgcacag cagaagacat gcttaatcct  
 10141 aactatgaag atctgctcat tgcctatcc aacctatgct ttcttggtca ggctggcaat  
 10201 gttcaacttc gtgttattgg ccattctatg caaaattgtc tgcctaggct taaagttgat  
 10261 acttctaacc ctaagacacc caagtataaa ttgttccgta tccaacctgg tcaaacattt  
 10321 tcagttcttag catgtacaaa tgggtcacca tctggtgttt atcagtgctg catgagacct  
 10381 aatcatacca ttaaagggttc ttctcttaat ggatcatgtg gtagtggttg ttttaacatt  
 10441 gattatgatt gcgtgtcttt ctgctatagt catcatatgg agcttccaac aggagtacac  
 10501 gctggtagtg acttagaagg taaattctat ggtccatttg tgcacagaca aactgcacag  
 10561 gctgcaggtg cagacacaac cataacatta aatgtttttg catggctgta tgcgtgctgt  
 10621 atcaatggtg ataggtggtt tcttaataga ttaccaccta ctttgaatga ctttaacctt  
 10681 gtggcaatga agtacaaacta tgaaccttgg acacaagatc atgttgacat attgggacct  
 10741 ctctctgctc aaacaggaaat tgcctgtcta gatatgtgtg ctgcttttga agagctgctg  
 10801 cagaattgta tgaatggctg tactatcctt ggtagcacta tttagaaga tgagtttaca  
 10861 ccatttgatg ttgttagaca atgctctggt gttaccttcc aaggtaaagt caagaaaaatt  
 10921 gttaaaggga ctcatacttg gatgctttta actttcttga atgctttctt gccatttact  
 10981 caaagtacac agtgggtcact gtttttcttt gtttacgaga atgctttctt gccatttact  
 11041 ttggtgatta tggcaattgc tgcattgtgt atgctgcttg ttaagcataa gcacgcattc  
 11101 cctgtagctg ttctgttacc ttctcttgca acagttgctt actttaatat ggtctacatg  
 11161 ggttataggc ttaaggattg tgttatgtat gcttcagctt tggctgacac tagcttgtct  
 11221 acagctcgca ctggttatga tgaatgctgt agctgtgctt ggacactgat taattctcatg  
 11281 ttagttattt ctgtaacctc ctatggtaat gctttagatc agctattatc catglgggccc  
 11401 agagctatag tgtttgtgtg taactattct ggtgtcgtta cgactatcat gtttttagct  
 11461 ttacagtgtg tcatgcttgt ttattgttcc tttagctatt taccattgtt tattattacc tggcaacacc  
 11521 cttttctggt tactcaaccc ttacttcagg caggggcttt tgcctcctaa gagtagtatt  
 11581 ttacacacaag aatttaggta tatgaacctc caggggcttt tgcctcctaa gagtagtatt  
 11701 gctgctttca agcttaacat taagtgtgtg ggtattggag gtaaacctat tatcaaggtt  
 11761 gtagctgtac agtctaaaaat gtctgacgta aagtgcacat ctgtggtact gctctcggtt  
 11821 cttcaacaac ttagagtaga gtcatctctt aaattgtggg cacaatgtgt acaactccac  
 11881 aatgatattc ttcttgcaaa agacacaact gaagctttcg agaagatggt ttctcttttg  
 11941 tctgttttgc tatccatgca ggtgctgta gacattaata ggttgtgca ggaatgctc  
 12001 gataaccgtg ctactcttca ggctattgct tcagaattta gttctttacc atcatatgcc

Fig 10. Cont.

12061 gcttatgcca ctgcccagga ggccctatgag caggctgtag ctaatgggtga ttctgaagtc  
 12121 gttctcaaaa agttaaagaa atctttgaat gtggtataat ctgagtttga ccgtgatgct  
 12181 gccatgcaac gcaagtttga aaagatggca gatcaggcta tgacccaaat gtacaaaacag  
 12241 gcaagatctg aggacaagag ggcaaaagta actagtgtca tgcaaaacaat gctcttcaat  
 12301 atgtcttaga agcttgataa tgaatgcact aacaacatta tcaacaatgc gcgtgatggt  
 12361 tgtgttccac tcaacatcat accattgact accatgcagca aactaatggt tgttgtccct  
 12421 gattatggta ccaacaagaa cacttggat ggtaacacct ttacatatgc atctgcactc  
 12481 tgggaaatcc agcaagtgtg tgaatggat agcaagattg ttcaacttag tgaatttaac  
 12541 atggacaatt caccaaaatt ggcttggcct cttattgtta cagctctaaag agccaaactca  
 12601 gctgttaaac tacagaataa tgaactgagt ccagtagcac tagcacagat gtctgtgcyg  
 12661 gctggtagca cacaacagc ttgtactgat gacaatgcac ttgcctacta taacaattog  
 12721 aaggagggtg ggtttgtgct ggcattacta tcagaccacc agatatctca atgggctaga  
 12781 ttccctaaga gtgatgtgac aggtacaatt tacacagaac tggaaaccacc ttgtaggttt  
 12841 gttacagaca caccaaaagg gctaaaagtg aaatacttgt acttcatcaa aggcctaaac  
 12901 aacctaataa gaggtatggt gctgggcagt ttagctgcta cagtacgtct tcaggctgga  
 12961 aatgctacag aagtacctgc caattcaact gtgctttcct tctgtgcttt tgcagtagac  
 13021 cctgctaag catataagga ttacctagca agtggaggac aaccaatcac caactgtgtg  
 13081 aagatgttgt gtacacacac tggtagagga caggcaatta ctgtaacacc agaagctaac  
 13141 atggaccaag agtcccttgg tgggtcttca tgttgtctgt attgtagatg ccacattgac  
 13201 catccaatc ctaaggatt ctgtgacttg aaaggttaagt acgtccaaat acctaccact  
 13261 ttgttaaatg acccagtggg ttttacactt agaaacacag tctgtaccgt ctgcgggaatg  
 13321 tggaaaggtt atggcttag ttgtgacctt cccgcgaac ccttgatgca ccgtgcggga  
 13381 gcataacgt ttttaaacgg gtttgcggtg taagtgcgc ccttcttaca ccgtgcggga  
 13441 caggactag tactgatgtc gtctacaggg cttttgatat ttacaacgaa aaagtgtgctg  
 13501 gttttgcaaa gttccataaa actaatgct gtgccttcca ggagaaggat gaggaaggca  
 13561 atttattaga ctcttacttt gtagttaaga ggcatactat gtctaactac caacatgaag  
 13621 agactattta taacttgggt aaagattgtc cagcggttgc tgtccatgac tttttcaagt  
 13681 ttagagtaga tggtagacatg gtaccacata tatcacgtca ggtctaaact aaatacacaa  
 13741 tggctgattt agtctatgct ctacgtcatt ttgatgaggg taattgtgat acattaaaag  
 13801 aaatactcgt cacatacaat tgcgtgtatg atgattatt caataagaag gattgggtatg  
 13861 actcgtaga gaatcctgac atcttaccgg tatatgctaa cttaggtgag cgtgtacgcc  
 13921 aatcattatt aaagactgta caattctgcy atgctatgcy tgatgcaggc attgtaggcy  
 13981 tactgacatt agataatcag gatcttaatg ggaactggtg cgatttcggt gatttcgtac  
 14041 aagtagcacc aggtgcgga gttcctattg tggattcata ttactcattg ctgatgccc  
 14101 tctcacttt gactaggga ttggctgctg agtcccatat ggtatgctgag ctgcgcaaac  
 14161 cacttattaa gtgggatttg ctgaaatattg attttacgga agagagactt tgtctcttcg  
 14221 ataggtattt taaatatttg gaccagacat accatcccaa ttgtattaac tgtttggatg  
 14281 accgtgtgat ccttcatttg gcaaacatta atgtgtatt tctactgtg ttccaccta  
 14341 caagttttgg accactagta gaaaaatat ttgtagatgg tgttcccttt gttgtttcaa  
 14401 ctggatacca ttttcgtgag ttaggagtcg tacataatca ggatgtaac catcatagct  
 14461 cgggtctcag tttcaaggaa cttttagttg atgctgtatg tccagctatg catgcagctt

5

10

15

20

25

30

35

40

Fig 10. Cont.

14521 ctggcaattt attgtagat aaacgcacta catgcttttc agtagtgca ctaacaaaca  
 14581 atgttgcttt tcaaacgtgc aaacccggta attttaataa agacttttat gactttgctg  
 14641 tgtotaaagg ttctttaag gaaggaagt ctgttgaaat aaacacttc tctttgctc  
 14701 aggatggcaa cgtgctatc agtgattatg actattatcg ttataatctg ccaacaatgt  
 14761 gtgatatacag acaactccta ttctgatttg aagtgttga taaatacttt gattgttacg  
 14821 atggtggctg tattaatgcc aaccaatgaa tcgttaacaa tctggataaa tcagctggtt  
 14881 tcccatthaa taaatgggt aagcctagac ttattatga ctcaatgagt tatgaggatc  
 14941 aagatgcact ttctgcgtat actaagcgta atgtcatccc tactataact caaatgaatc  
 15001 ttaagtatgc cattagtgc aagaatagag ctgcacccgt agctggtgc tctatctga  
 15061 gtactatgac aaatagacag ttatcatcaga aattattgaa gtcaatagcc gccactagag  
 15121 gagctactgt ggtaattgga acaagcaagt ttacgggttg ctggcataat atgttaaaaa  
 15181 ctgtttacag tgatgtagaa actccacacc ttatgggttg ggattatcca aaatgtgaca  
 15241 gagccatgcc taacatgctt aggataatgg cctctcttgt tctgtctgcg aaacataaca  
 15301 ctgtgctgtaa cttatcacac cgtttctaca ggttagctaa cgagtgtgcg caagtattaa  
 15361 gtgagatggt catgtgtggc ggctcactat atgttaaac aggtgaaaca tcatccggtg  
 15421 atgctacaac tgcctatgct aatagtgtct ttaacatttg tcaagctgtt acagccaatg  
 15481 taaatgcact tctttcaact gatggttaata agatagctga caagtatgtc cgcaatctac  
 15541 aacacaggct ctatgagtgt ctctatagaa atagggatgt tgcgtgatg  
 15601 agttttacgc ttacctgctt aaacatttct ceatgatgat tctttctgat gatgccgttg  
 15661 tgtgctataa cagtaactat gcggctcaag gtttagtagc tagcattaaag aactttaagg  
 15721 cagttcttta ttatcaaaat aatgtgttca tgcctgagcc aaaatgttgg actgagactg  
 15781 accttactaa aggacctcac gaattttgct cacagcatac atgtctagtt aaacaaggag  
 15841 atgattacgt gtacctgcct taccagatc catcaagaat attaggcgca ggctgttttg  
 15901 tcgatgatat tgtcaaaaca gatggtacac ttatgattga aaggttcgtg tcactggcta  
 15961 ttgatgctta ccacttaca aaacatccta atcaggagta tgcgtgatgt tttcaattgt  
 16021 attacaata cattagaaag ttacatgatg agcttactgg ccacatgttg gacatgtatt  
 16081 ccgtaatgct actaatgat aacacctcac ggtactggga acctgagttt tatgaggcta  
 16141 tgtacacacc acatacagtc ttgcaggctg taggtgcttg tgtattgtgc aattcacaga  
 16201 ctacacttgc ttgcgggtgc tgtattagga gaccttccct atgttgcaag tgcgtctatg  
 16261 accatgtcat ttcaacata cacaatttag tgtgtctgt taatccctat gtttgcaatg  
 16321 cccagggttg tgatgtcaat gatgtgacac aactgtatct agggaggtatg agctattatt  
 16381 gcaagtaca taagctccc attagtcttc cattatgtgc ttcaatgctg gtttttggtt  
 16441 tatacaaaaa cacatgtgta ggcagtgaca atgtcactga cttcaatgct atagcaacat  
 16501 gtgattggac taatgctggc gattacatac ttgccaacac ttgtactgag agactcaagc  
 16561 ttttcgacgc agaacgctc aaagccactg aggaacatt taagctgtca tatggtattg  
 16621 ccactgtaag cgaagtactc tctgacagag aattgcactc ttcattggag gttggaacac  
 16681 tagaacacc attgaacaga aactatgtct ttactggta ccggtgaact aaaaatagta  
 16741 atgacagat tggaggtac acctttgaaa aaggtgacta tgggtgatgct gttgtgtaca  
 16801 gaggtactac gacatacaag ttgaatgttg gtgattactc tgtgttgaca tctcacactg  
 16861 taatgccaact tagtgcaact actctagctc cacaaggaca ctatgtgaga attactggct  
 16921 tgtacccaac actcaacatc tcagatgagt tttctagcaa tgttgcaat tatcaaaagg

5

10

15

20

25

30

35

40

Fig. 10. Cont.

16981 tcggcatgca aagtactct acactccaag gaccacctgg tactggttaag agtcattttg  
 17041 ccatcgact tgcctctat taccatctg ctgcgcatagt gtatacggca tgcctctcatg  
 17101 cagctgttga tgcctatagt gaaaaggcat taaaatttt gcccatagat aaatgtagta  
 17161 gaacataacc tgcggtgctg cgcgtagagt gttttgataa attcaaatg aattcaaac  
 17221 tagaacagta tgtttctgc actgtaaatg cattgccaga acaactgct gacattgtag  
 17281 tctttgatga aatctctatg gctactaat atgacttgag tgtgtcaat gctagacttc  
 17341 gtgcaaaaaca ctacgtctat attggcgatc ctgctcaatt accagcccc cgcacattgc  
 17401 tgactaaaagg cacactagaa ccagaatatt ttaattcagt gtgcagactt atgaaaacaa  
 17461 taggtccaga catgttcctt ggaacttgc gccgttgcct tgcgaaatt gttgacactg  
 17521 tgagtgcctt agtttatgac aataagctaa aagcacacaa ggataagtca gctcaatgct  
 17581 tcaaaatggt ctacaaagg tgaagagaa tttcttacac atgatgttc atctgcaatc aacagacctc  
 17641 aaataggcgt tgtaagagaa aacgctgtag cttcaaaaat cttaggattg cctacgcaga  
 17701 totcacctta taattcacag aacgctgtag cttcaaaaat cttaggattg cctacgcaga  
 17761 ctgttgattc atcacagggt tctgaatatg actatgtcat attcacaaa actactgaaa  
 17821 cagcacactc ttgtaatgtc aacgccttca atgtggctat cacaaggga aaaattggca  
 17881 ttttgcgcat aatgtctgat agagatcttt atgacaaact gcaatttaca agtctagaaa  
 17941 taccacgtcg caatgtggct acattacaag cagaaaatgt aactggactt tttaggact  
 18001 gtagtaagat cattaactgt cttcatccta cacaggcacc tacacacctc agcgttgata  
 18061 taaagttcaa gactgaagga ttatgtgttg acataccagg cataccaag gacatgacct  
 18121 accgtagact catctctatg atgggtttca aaatgaatta ccaagtcaat ggttacocct  
 18181 atagtgttat caccgcgaa gaagctatct gcacgttctg tgcgtggatt ggccttgatg  
 18241 tagaggctcg tcatgcaact agagatgctg tgggtactaa cctacctc cagctaggat  
 18301 ttctacagg tgttaactta gtactgtac cgactggtta tgtgacact gaaaataaca  
 18361 cagaattcac cagagttaat gcaaaacctc caccagggtg ccagtttaaa catcttatac  
 18421 cactcatgta taaaggcttg cctgggaatg tagtgcgtat taagatagta caaatgctca  
 18481 gtgataacat gaaaggattg tcagacagag tgcgtgtcgt cctttgggcg catggctttg  
 18541 agcttacatc aatgaagtac ttgtcaaga ttggacctga aagaacgtgt tgcctgtgtg  
 18601 acaaacgtgc aacttgcttt tctacttcat cagatactta tgcctgctgg aatcattctg  
 18661 tgggttttga ctatgtctat aaccattta tgattgatgt tcagcagtgg ggccttaacgg  
 18721 gtaacctca gattaacct gagtaacct gccagggtaca tggaaatgca catgtggcta  
 18781 gttgtgatgc tatcatgact agatgtttag cagtcctatga gtgctttgtt aagcgcgttg  
 18841 attggtctgt tgaataacct attataggag atgaactgag ggttaattct gcttgcagaa  
 18901 aagtaacaaca catggtttgt aagctgcat tgcctgtcga taagtttcca gttcttcattg  
 18961 acattggaaa tccaaaggct atcaagtgtg tgcctcaggc tgaagttaga tggaaattct  
 19021 acgatgtctca gcatgtagt gacaaagttt tttgtttgtt ttggaattgt aacgttgatc  
 19081 ctacacatca cgataaattc actgatggtg ttgacacaa agtcttgtca aacttgaact  
 19141 ttacccagc caatgcaatt gtgtgtaggt tgaataagca tgcattccac actccagctt  
 19201 taccagcgtg tgatgggtgt agtttgtatg tgaataagca tgcattccac actccagctt  
 19261 tcatgaaaag tgcatttact aatttaagc aattgcctt cttttactat tctgatagtc  
 19321 ctgtgtgagtc tcatggcaaa caagtgtgt cggatattga ttatgttcca ctcaaatctg  
 19381 ctacgtgtat tacacgatgc aatttaggtg gtgctgtttg cagacaccat gcaaatgagt



Fig. 10. Cont.

19441 accgacagta cttggatgca tataatatga tgattttgc tgattttagc ctatggattt  
 19501 acaacaatt tgatacttat aacctgtgga atacatttac caggttacag agtttagaaa  
 19561 atgtggctta taatgttgtt aataaggac actttgatgg tattgatgg gagatctttg  
 19621 ttccatcat taataatgct gttacacaa agtagatgg cattgcttaag cgtaacatta  
 19681 aaataagac aacacttctt gttatgttg catttgagct ttgggctaaag gtaataactg  
 19741 aaccatggcc agagattaag atactcaata atttgggtgt tgatatcgt gtaataactg  
 19801 taactctggga ctacaaaaga gaagcccccag cacattgatac tacaataggt gtctgcacaa  
 19861 tgactgacat tgcacaagaa cctactgaga gtgcttgttc ttacattact gtcttgtttg  
 19921 atggtagagt ggaaggacag gtagaccttt ttagaacgc cgtaatgtt gtttaataa  
 19981 cagaaggttc agtcaaaagt ctaaacacctt caaaggacc agcacaagt agcgtcaatg  
 20041 gagtacatt aattggagaa tcaagtaaaa cacagtttaa ctactttaag aaagtagacg  
 20101 gcattattca acagttgcct gaaacctact ttactcagag cagagactta gaggtattta  
 20161 agccagatc acaaatggaa actgactttc tccagctcgc tatggatgaa ttcatacagc  
 20221 gatataagct cgagggtctat gccctgaac acatcgttta tggagatttc agtcatggac  
 20281 aacttggcgg tcttcattta atgataggct tagccaagcg ctcaacaagt tcaccactta  
 20341 aattagagga ttttatccct atggacagca cagtgaaaaa ttacttcata acagatggcg  
 20401 aaacaggttc atcaaaatgt gtgtgttctg tttcaaaagt tttacttgat gactttgtcg  
 20461 agataataaa gtcacaagat ttgtcagtga tttcaaaagt ggtcaaggtt acaattgact  
 20521 atgctgaaat ttcatttcag ctttgggtga aggatggaca tgttgaacc ttctaccaa  
 20581 aactacaagc aagtcaagcg tggcaaccag gtgttgcgat gctaaacttg tacaagatgc  
 20641 aaagaatgct tcttgaagaag tgtgaccttc agaattatgg tgaaaatgct gttataccaa  
 20701 aggaataaat gatgaatgct gcaaatgata ctaactgtg tcaatactta aatacaccta  
 20761 ctttagctgt accctacaac atgagagtta ttcaacttgg tctggctct gataaaggag  
 20821 ttgcaccagg tacagctgtg ctacagacaa ggttgccaac tggcacacta cttgtcgatt  
 20881 cagatcttaa tgacttcgtc tccgacgcag attctacttt aattggagac tgtgcaacag  
 20941 tacatacggc taataaatgg gacctatta tttagcgatat gtatgacctt aggaaccaac  
 21001 atgtgacaaa agacaaatgg agccctgggt ggttctatag ctgtaaaagt aacagagcat tcttgggaatg  
 21061 agcaaaaact agcccttctg ggcctattct catggtggac agcttttgtt acaaatgtaa  
 21121 ctgaccttta caagcttatg ggcctattct catggtggac agcttttgtt acaaatgtaa  
 21181 atgcatcatc atcggaagca tttttaattg ggcctaaact acattttctg gaggaacaca aatcctatcc  
 21241 aaattgatgg ctataccatg catgctaact tttagacatga gcaaatcttc tcttaaatga agaggaaactg  
 21301 agttgtcttc ctattcactc ttttaaggag aatcaaatca atgatatgat ttattctctt ctggaataag  
 21361 ctgtaaatgc tcttaaggag aatcaaatca atgatatgat ttattctctt ctggaataag  
 21421 gtaggcttat cattagagaa aacaacagag ttgtggtttc aagtgatatt cttgttaaca  
 21481 actaaacgaa catgtttatt tctttattat tcttactctt cactagtgggt agtgacattg  
 21541 accggtgac cactttttgat gatgttcaag ctctcaatta cactcaacat acttcatcta  
 21601 ttgaggggggt ttactatcct tctaatttga ttagatcaga cactctttat ttaactcagg  
 21661 atttatttct tccattttat tctaatttga ttagatcaga cactctttat ttaactcagg  
 21721 gcaacctgt catacctttt aaggatggtta tttattttgc tgcacagag aaatcaaatg  
 21781 ttgtcgtggg ttgggttttt ggttctacca tgaacaacaa gtcaacgtcg gtgatttata  
 21841 ttaacaattc tactaatgtt gttatacagc catgtaactt tgaattgtgt gacaacctt

Fig. 10. Cont.

21901 tctttgctgt ttctaaaccc atgggtacac agacacatac tatgatattc gataatgcat  
 21961 ttaattgcac ttctgagtag atatctgatg ccttttcgct tgatgtttca gaaaagtcag  
 22021 gtaattttaa acacttacga gagtttgtgt ttaaaaataa agatgggttt ctcctatgtt  
 22081 atagggtcta tcaacctata gatgtagtgc gtgacttacc ttctgggttt aacactttga  
 22141 aacctatttt taagtgtgct cttggtatta acattacaaa ttttagagcc attcttacag  
 22201 cttttccacc tgcacaagac atttggggca cgtcagctgc agcctatttt gttggctatt  
 22261 ttaagccaac tacattttat ctcaagtatg atgaaaatgg tacaatcaca gatgctgttg  
 22321 attgttctca aaatccactt gctgaactca aatgctctgt taagagcttt gagattgaca  
 22381 aggaattta ccagacctct aatttcaggg ttgttccctc aggagatgtt gtgagattcc  
 22441 ctaatattac aaacttgtgt ccttttggag aggtttttaa tgcactaaa ttccctcttg  
 22501 tctatgcatt ggagagaaaa aaaatttcta attgtgttgc tgattactct gtgctctaca  
 22561 actcaacatt tttttcaacc tttaagttcc ttgagttcct atggcgtttc tggcactaag ttgaatgac  
 22621 ttgctctc caatgtctat ttcagattct cagagttcct ttgtagtcaa gggagatgat gtaagacaaa  
 22681 tagcgcagg acaaaactgt gttatgtcgt attataatta taaattgcca gatgattca  
 22741 tgggttgtgt ccttgccttg aatactagga acattgatgc tactcaact ggtaattata  
 22801 attataaata taggtatctt agacatggca agcttaggcc ctttgagaga gacatatcta  
 22861 atgtgcctt ctccctgat ggcaaacctt gcacccacc tgcctttaat tgttattggc  
 22921 cattaaatga ttatggtttt tacaccacta ctggcattgg ctaccaacct tacagagtgg  
 22981 tagtactttc ttttgaactt ttaaatgcac cggccacggg ttgtggacca aatttatcca  
 23041 ctgaccttat taagaaccag tgtgtcaatt ttaattttaa tggactcact ggtactgggtg  
 23101 tgttaactcc ttcttcaag agatttcaac catttcaaca atttggcgt gatgtttctg  
 23161 atttcaatga ttccgttcca gatcctaaaa catctgaaat attagacatt tcaccttggg  
 23221 cttttggggg tgaagtgtta attacacctg gaacaaatgc tcaatctgaa ttgctgttcc  
 23281 tatatcaaga tgttaactgc actgatgttt ctacagcaat tcaatgcagat caactcacac  
 23341 cagcttggcg catatatctt actggaacaa atgtattcca gactcaagca ggctgtctta  
 23401 taggagctga gcatgtcgac acttcttatg agtgcgacat tccatttggg gctggcattt  
 23461 gtgctagtta ccatacagtt tctttattac gtagtactag ccaaaaaatct atttggcctt  
 23521 atactatgtc tttaagtgtc gatagtcca ttgcttactc taataacacc attgtatatac  
 23581 ctactaactt ttcaattagc attactacag aagtaaatgcc tgtttctatg gctaaaaact  
 23641 ccgtagattg taatatgtac atctgcggag attctactga atgtgctaatt ttgcttctcc  
 23701 aatatgtag cttttgcaca caactaaatc gtgcactctc aggtattgct gctgaacagg  
 23761 atcgcaacac acgtgaagtg ttcgctcaag tcaaacaaat gtacaaaacc ccaactttga  
 23821 ataatatttg tggtttttaatt ttttcaaaa tattacctga cctctctaaag ccaactaaga  
 23881 ggtcctttat tgaggacttg ctctttaata aggtgacact cgctgatgct ggcctcatga  
 23941 agcaatatgg cgaatgccta ggtgatatta atgctagaga tctcatttgt ggcgagaagt  
 24001 tcaatggact tacagtgttg ccacctcgc tcaactgata tatgattgct gctacactg  
 24061 ctgctctagt tagtggctact gccactgctg gatggacatt tgggtcgtggc gctgctcttc  
 24121 aaataccttt tgcctatgcaa atggcatata ggttcaattg cattggagtt acccaaatg  
 24181 ttctctatga gaaccaaaaa caaatgcoca accaatttta caagggcatt agtcaaatc  
 24241 aagaatcact tacaacaaca tcaactgcat tgggcaagct gcaagacgtt gtttaaccaga  
 24301 atgtcaagc attaaacaca ctgtttaaac aacttagctc taattttggt gcaatttcaa

5

10

15

20

25

30

35

40

Fig. 10. Cont.

24361 gtgtgctaaa tgatatcctt tgcgacttg ataaagtctga ggcggaggta caaattgaca  
 24421 ggtaattac aggcagactt caaagccttc aaacctatgt aacacacaa ctaatcaggg  
 24481 ctgtcgaat cagggtcttc gctaactctg ctgtactaa atgtctgag tgtgttcttg  
 24541 gacaatcaaa aagagttgac ttgttgaa agggctacca cttatgtcc ttcccacaag  
 24601 cagcccgca tgggtgtgtc ttctacatg tcacgtatgt gccatccag gagaggaact  
 24661 tcaccacag gccagcaat tgcacatg gcaagcata ctccctcgt gaagggtgtt  
 24721 ttgtgtttaa tggcacttct tggtttatta cacagaggaa cttctttct ccacaataa  
 24781 ttactacaga caatacatt gtctcaggaa attgtgatgt cgttatggc atcataaaca  
 24841 acacagttta tgatcctctg caacctgagc ttgactcatt caaagaagag ctggacaagt  
 24901 acttcaaaaa tcatacatca ccagatgttg atcttggcga cattcaggc attaacgctt  
 24961 ctgtcgtcaa cattcaaaa gaaattgacc gctcaatga ggtcgctaaa aatttaaatg  
 25021 atcacactcat tgaccttcaa gaattgggaa aatatgagca atatatataa tggccttggt  
 25081 atgtttggct cggcttcat gctggactaa ttgccatcgt catggttaca atcttgcttt  
 25141 gtgcacgac tagttgttc agttgcctca aggtgcatg ctcttggtgt tcttgctgca  
 25201 agtttgatga ggaatgactct gagccagttc tcaagggtgt caaattacat tacacataaa  
 25261 cgaacttatg gatttgttta tgagatttt tactcttaga tcaattactg cacagccagt  
 25321 aaaaattgac aatgcttctc ctgcaagtac tgttcactgt acagcaacga tacogctaca  
 25381 agcctcactc cctttcggat ggcttggtat tggcgttgc tttcttgctg ttttcaagag  
 25441 cgctaccaaa ataatggcg tcaataaaag atggcagcta gcccttata aggccttcca  
 25501 gtccatttgc aatttactgc tgcattttgt taccatctat tccatcttt tgcctgtcgc  
 25561 tgcaggatg gagcgcaat tttgtacct ctatgacctg atatatattc tacaatgcac  
 25621 caacgcattg agaattatta actttgtttg gcttgttgg agtgcaaat ccaagaaccc  
 25681 attactttat gatgccaaat actttgtttg ctggcacaca cataactatg actactgtat  
 25741 accataaac agtgcacag atacaattgt cgttactgaa ggtgacggca ttcaaacacc  
 25801 aaactcaaa gaagactacc aaattggtg ttattctgag gataggcact caggtgttaa  
 25861 agactatgc gtgtacatg gctatttcac cgaagtttac taccagcttg agtctacaca  
 25921 aattactaca gacactggtg ttgaaaatgc tacattcttc atctttaaca agcttgtaa  
 25981 agaccacag aatgtgcaaa tacacacaaat cgacggctct tcaaggagtgt ctaatccagc  
 26041 aatggatcca atttatgatg agccgacgac gactactagc gtgcctttgt aagcacaaga  
 26101 agtgagttac gaacttatgt actcattcgt ttcggaagaa acaggtacgt taatagttaa  
 26161 tagcgtactt ctttttcttg ctttctgtgt attcttgcta gtcacactag ccactcctac  
 26221 tgcgtcttga ttgtgtcgt actgctgcaa tattgttaac ttgagttag taaaaccaac  
 26281 ggtttaactg tactcgcgtg ttaaaatct gaactcttct gaaggagtcc ctgatcttct  
 26341 ggtctaaacg aactaaactat tattattatt cgttttgaa ctttaacatt gcttatcatg  
 26401 gcagacaacg gtactattac cgttgaggag cttaaacac tccctggaaca atggaacct  
 26461 gtaaatggtt tctattcct agcctggtt atgttactac aatttgccta ttctaactcg  
 26521 aacagggttt tgtacataat aaagcttgtt ttccctggc tcttggtggc agtaaacatt  
 26581 gcttggtttt tgatcgtgc ttgtacaga attaatggg tgaatggcg gattgcgatt  
 26641 gcaatggctt gtatttagg cttgatgttg cttagctact tggttgcttc ctcaaggctg  
 26701 ttgtctcgta cccgctcaat gtggtcattc aaccagaaa caaacattct tctcaatgtg  
 26761 cctctccggg ggacaattgt gaccagacg ctcatggaaa gtgaacttgt cattggtgct

5

10

15

20

25

30

35

40

Fig. 10. Cont.

26821 gtgatcattc gtggtcactt gogaatggcc ggacactccc tagggcgctg tgacattaa  
 26881 gacctgcaa aagagatcac tgtgggtaca tcacgaacgc ttctctatta caaattagga  
 26941 gcgtcgagc gtgtaggcac tgattcaggt tttgtgcac aaacgccta ccgtattgga  
 27001 aactataaat taaatacaga ccacgccggt agcaacgaca atattgcttt gctagtacag  
 27061 taagtacaa cagatgtttc atctgttga cttccaggtt acaatagcag agatattgat  
 27121 tatcattatg aggaacttca ggaattgctat ttggaatttt gacgttataa taagtccaat  
 27181 agtgagacaa ttatttaagc ctctaactaa gaagaattat tcggagttag atgatgaaga  
 27241 acctatggag ttagattatc cataaaacga acatgaaaat tattctcttc ctgacattga  
 27301 ttgtatttac atcttgcgag ctatatcac atcaggagtg tttacacattt caccctcttg  
 27361 tactaaaga accttgccca tcagggaacat acgagggcaa ttacacactt tgcctttgct  
 27421 ctgacaataa atttgcacta acttgcacta gacacacatt tgcctttgct tgtgtgacg  
 27481 gtactcgaca taactatcag ctgctgtgcaa gacagtttc accaaaactt ttcctcagac  
 27541 aagaggaggt tcaacaagag ctctactcgc cactttttct cattgttgc tctcagac  
 27601 ttttaatact ttgcttcacc attaagagaa agacagaatg attccttgtt ttaataatgc  
 27661 ctctatttgc ctgaaatcc aggttctgct ctttctgctt agaaccttgc accaaagtct  
 27721 ttggttttca ctgaaatcc aggttctgct ctttctgctt agaaccttgc accaaagtct  
 27781 gaaacttttc attgttttga cttgtatttc tctatgcagt tgcataatgc ctttaattga  
 27841 gcgctgtgca tctaataaac ctcatgtgct tgaagatcct tgaaggttac aactagtagg  
 27901 gtaatactta tagcactgct tggctttgtg ctctaggaaa ggttttacct ttctatagat  
 27961 ggacacacta ggttcaaaaca tgcacaccta atgttactat caactgtcaa gatccagctg  
 28021 gtggtgcgct tatagtagg ttgtgtgtacc ttcatgaag tcaccaaaact gctgcattta  
 28081 gagacgtact tgttgtttta aataaagcaa caaattaaaa tglctgataa tggaccccaa  
 28141 tcaaaccaac gtatgtcccc ccgcattaca ttgtgtggac ccacagattc aactgacaat  
 28201 aaccagaatg gaggacgcaa tggggcaagg ccaaaacagc gccgacccca aggtttaccc  
 28261 aataatactg cgtcttgggt cacagctctc actcagcatg gcaaggagga acttagattc  
 28321 cctcgaggcc agggcggttc aatcaacacc aatagtgttc cagatgacca aattggctac  
 28381 taccgaagag ctacccgacg agttcgtgtt ggtgacggca aatgaaaga gctcagcccc  
 28441 agatgttact tctattacct aggaactggc ccagaagctt cacttcccta cggcgctaac  
 28501 aaagaggcca tcgtatgggt tgcaactgag ggagccttga atacacccaa agaccacatt  
 28561 ggacaccgca atcctaataa caatgctgcc accgtgctac aacttctca aggaacaaca  
 28621 ttgccaacag gcttctacgc agagggaagc agaggcgga gtcaagcctc ttctcgctcc  
 28681 tcatacgtg gtgcggttaa ttcaagaaat tcaactcctg gcagcagtag gggaatttct  
 28741 cctgtctgaa tggctagcgg aggtgtgtgaa actgcccctg cgctattgct gctagacaga  
 28801 ttgaccagc ttgagagcaa agtttctggt aaaggccaac aacaacaagg ccaactgtc  
 28861 actaagaat ctgctgctga ggcattcaaa agcctcgcc aaaaacgtac tgcacaaaa  
 28921 cagtaacaag acctaatcag aagggaact gattacaaac attggccgca aattgcacaa  
 28981 gggtacacac gtgcctctgc atcttctgga atgtcagca ttggcatgga agtcacacct  
 29041 ttgtctccaa ttgtctctgc tcatctgagc attaaattgg atgacaaaaga tccacaattc  
 29101 tcgggaacat ggctgactta tcatctgact gaacaagcac attgacgcat acaaacattt ccaccaaca  
 29161 aaagacaacg tcatactgct gaaacaagc gaaacaagc gatgaagctc agcctttgcc gcagagacaa  
 29221 gaggctaaaa aggcacaaaa gaaaagact gatgaagctc agcctttgcc gcagagacaa

5

10

15

20

25

30

35

40

Fig. 10. Cont.

29281 aagaagcagc ccaactgtgac tcttcttctcct ggggctgaca tggatgattt ctccagacaa  
29341 cttcaaaatt ccatgagtgg agtttctgct gattcaactc aggcataaac actcatgatg  
29401 accacacaag gcagatgggc tatgtaaacg ttctcgcaat tccgtttacg atacatagtc  
29461 tactcttgtg cagaatgaat tctcgtaact aaacagcaca agtaggttta gttaaacttta  
29521 atctcacata gcaatcttta atcaatgtgt aacattaggg aggaattgaa agagccacca  
29581 cattttcatc gaggccacgc ggagtcacgat cgagggtaca gtgaataatg ctaggagagag  
29641 ctgcctatat ggaagagccc taatgtgtaa aattaatttt agtagtgcta tccccatgtg  
29701 attttaatag ctctcttaga gaatgacaaa aaaaaaaaaa a

34/36

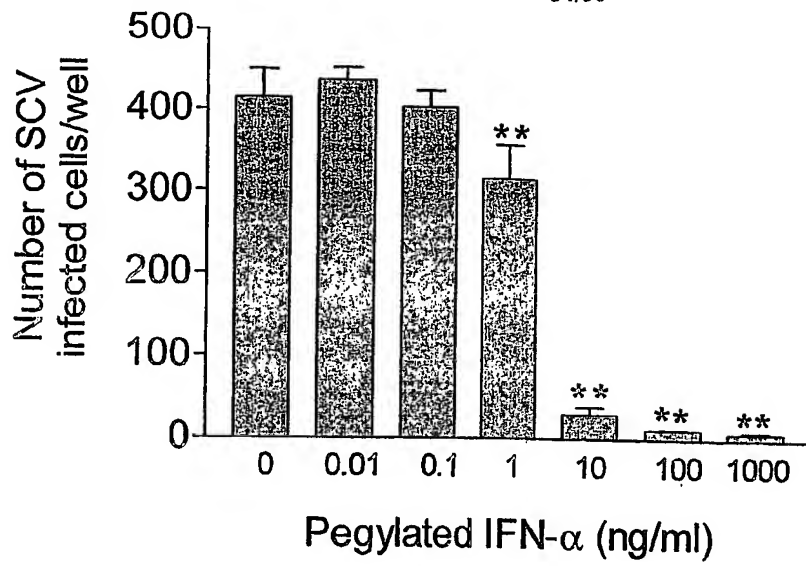
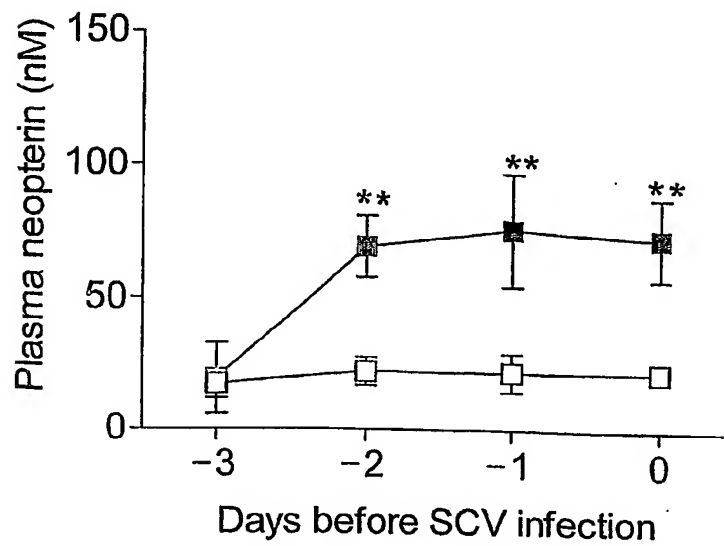
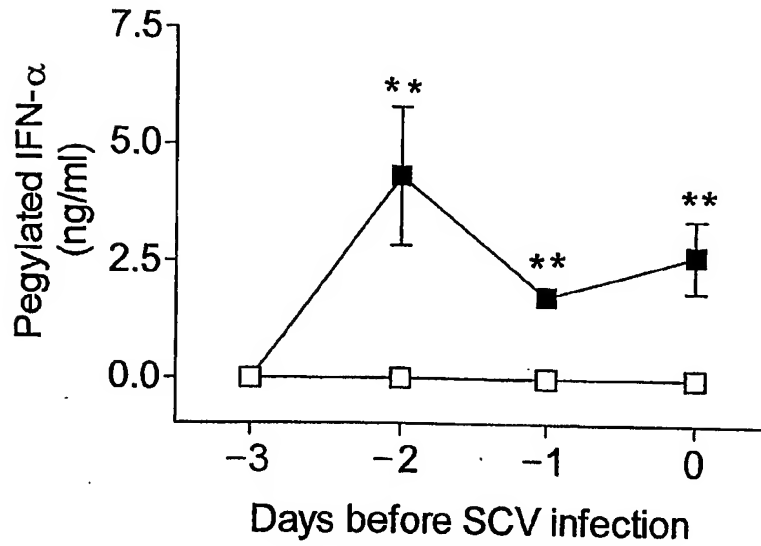


Fig 11



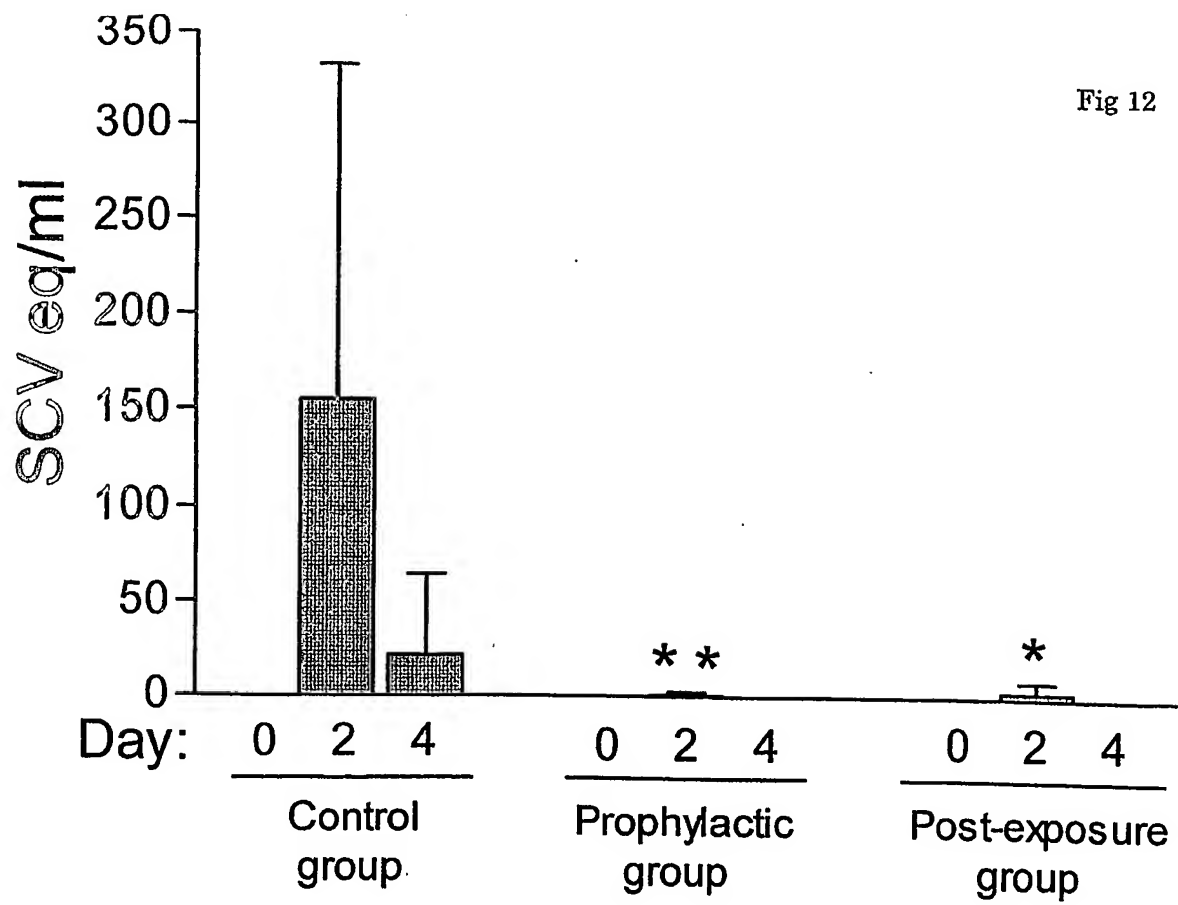
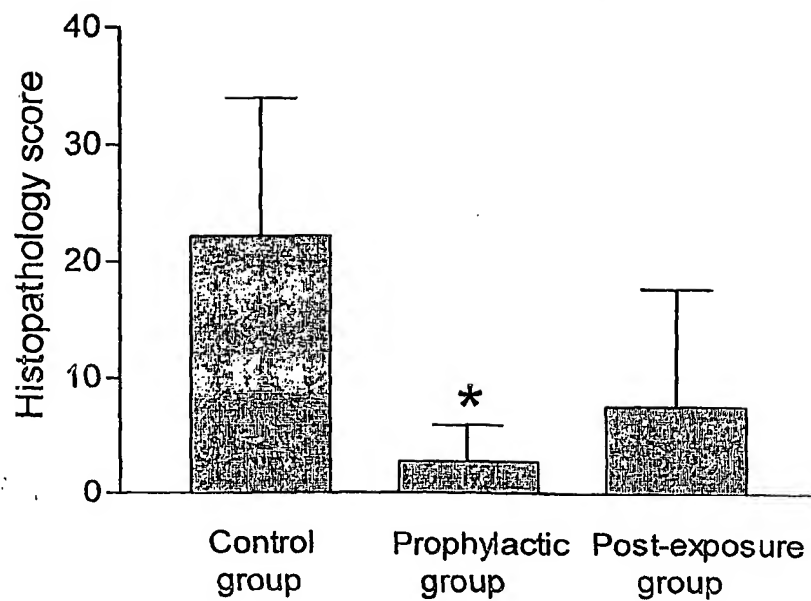
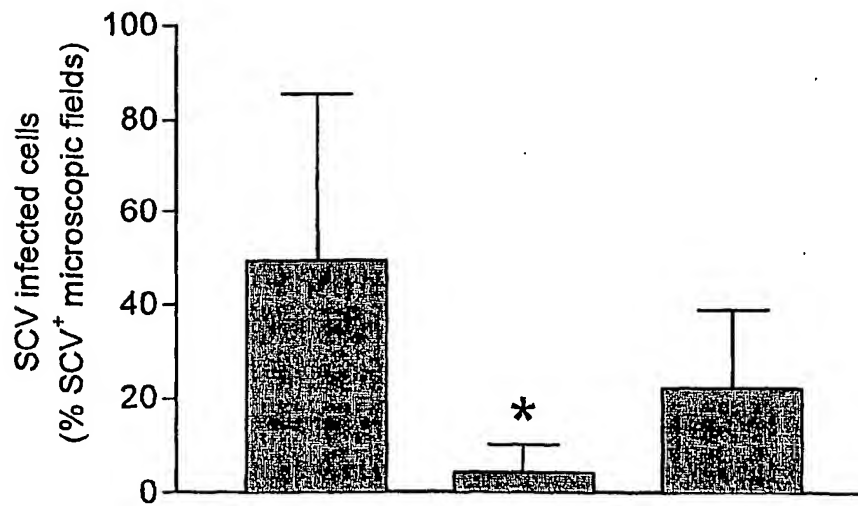
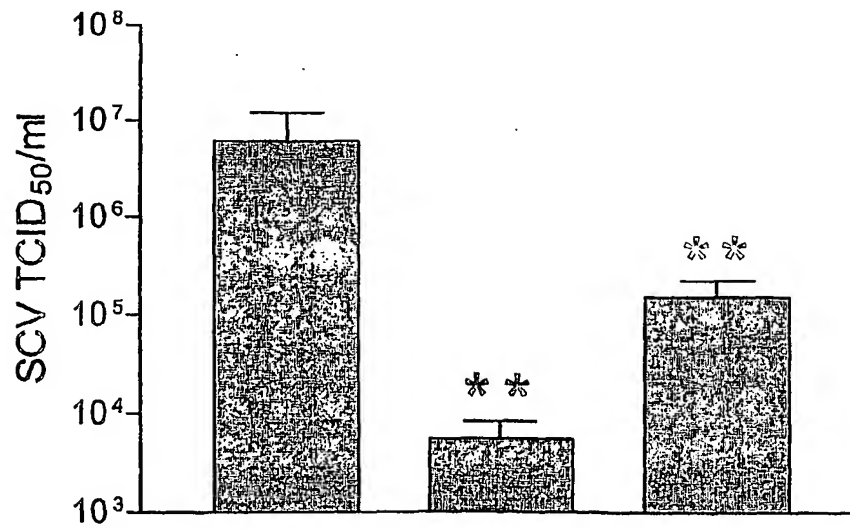


Fig 13





(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
21 October 2004 (21.10.2004)

PCT

(10) International Publication Number  
**WO 2004/089983 A3**

(51) International Patent Classification<sup>7</sup>: **C07K 14/165**,  
C12N 15/85, 5/10, C07K 16/10, C12Q 1/70, G01N  
33/569, A61K 39/215, 38/21, A61P 31/14

(74) Agent: WINCKELS, J., H., F.; Nieuwe Parklaan 97,  
NL-2587 BN Den Haag (NL).

(21) International Application Number:  
PCT/NL2004/000229

(81) Designated States (unless otherwise indicated, for every  
kind of national protection available): AE, AG, AL, AM,  
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,  
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,  
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,  
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,  
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,  
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,  
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,  
ZW.

(22) International Filing Date: 8 April 2004 (08.04.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
03076037.5 8 April 2003 (08.04.2003) EP  
03076110.0 14 April 2003 (14.04.2003) EP  
03077307.1 18 July 2003 (18.07.2003) EP

(84) Designated States (unless otherwise indicated, for every  
kind of regional protection available): ARIPO (BW, GH,  
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Euro-  
pean (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR,  
GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK,  
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,  
ML, MR, NE, SN, TD, TG).

(71) Applicant (for all designated States except US): VI-  
RONOVATIVE B.V. [NL/NL]; Burgemeester Oudlaan  
50, NL-3062 PA Rotterdam (NL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): HAAGMANS,  
Bartholomeus, Leonardus [NL/NL]; Doude van Troost-  
wijkstraat 3, NL-1391 EP Abcoude (NL). KUIKEN,  
Thijs [NL/NL]; Middenmolenlaan 149, NL-2807 EV  
Gouda (NL). FOUCHIER, Ronaldus, Adrianus, Marla  
[NL/NL]; Essenburgsingel 44a, NL-3021 AR Rotterdam  
(NL). OSTERHAUS, Albertus, Dominicus, Marcellin-  
us, Erasmus [NL/NL]; Dr. Breveestraat 16, NL-3981  
CH Bunnik (NL).

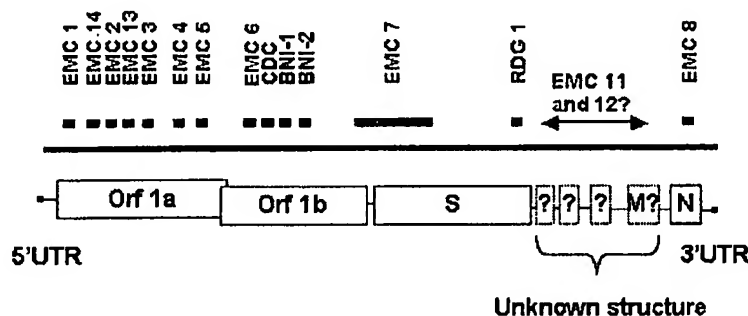
Published:

- with international search report
- before the expiration of the time limit for amending the  
claims and to be republished in the event of receipt of  
amendments

(88) Date of publication of the international search report:  
20 January 2005

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

(54) Title: SEVERE ACUTE RESPIRATORY SYNDROME (SARS) CAUSING CORONAVIRUS



(57) Abstract: The invention relates to the field of virology. The invention provides an isolated essentially mammalian positive-sense single stranded RNA virus (SARS) within the group of coronaviruses and components thereof. The effect of pegylated interferon-alpha on SARS infection has been analysed.

# INTERNATIONAL SEARCH REPORT

ational Application No

/NL2004/000229

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C07K14/165 C12N15/85 C12N5/10 C07K16/10 C12Q1/70  
G01N33/569 A61K39/215 A61K38/21 A61P31/14

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 A01K C07K C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, Sequence Search, BIOSIS, EMBASE

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>POUTANEN SUSAN M ET AL: "Identification of severe acute respiratory syndrome in Canada." NEW ENGLAND JOURNAL OF MEDICINE, vol. 348, no. 20, 31 March 2003 (2003-03-31), pages 1995-2005, XP002255633 ISSN: 0028-4793 page 2002, left-hand column, paragraph 3 - right-hand column, paragraph 3 ----- -/--</p>	1-27

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

27 October 2004

Date of mailing of the international search report

25/11/2004

Name and mailing address of the ISA  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Brouns, G

## INTERNATIONAL SEARCH REPORT

International Application No

/NL2004/000229

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CENTERS FOR DISEASE CONTROL AND PREVENTION: "Update: Outbreak of severe acute respiratory syndrome: Worldwide, 2003." MORBIDITY AND MORTALITY WEEKLY REPORT, vol. 52, no. 12, 28 March 2003 (2003-03-28), pages 241-248, XP002288825 ISSN: 0149-2195 page 243, paragraph 2 -----	1-27
P, X	DATABASE EMBL E.B.I. HINXTON, UK; 15 April 2003 (2003-04-15), S. JONES: "SARS Coronavirus TOR2, complete genome" XP002288827 retrieved from EMBL Database accession no. AY274119 EMC-1=nts 2064-2932;2=4820-5166;3=8480-8928;4=9954-10633;5=nts12194-13525;6=nts 13877-14648;7=nts 20665-23772;8=nts 28228-28590;11=611-2091;12=4325-4770;13=6338-7171;14=3579-4315;RDG1=24439-25090 abstract -----	1-27
P, X	SNIJDER E J ET AL: "Unique and Conserved Features of Genome and Proteome of SARS-coronavirus, an Early Split-off From the Coronavirus Group 2 Lineage" JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 331, no. 5, 29 August 2003 (2003-08-29), pages 991-1004, XP004447471 ISSN: 0022-2836 the whole document -----	1-27
P, X	PEIRIS J ET AL: "Coronavirus as a possible cause of severe acute respiratory syndrome" LANCET THE, LANCET LIMITED. LONDON, GB, vol. 361, no. 9366, 8 April 2003 (2003-04-08), pages 1319-1325, XP004421148 ISSN: 0140-6736 the whole document -----	1-27

-/--

## INTERNATIONAL SEARCH REPORT

International Application No

/NL2004/000229

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GERBERDING JULIE LOUISE: "Faster... but fast enough? Responding to the epidemic of severe acute respiratory syndrome." THE NEW ENGLAND JOURNAL OF MEDICINE. 15 MAY 2003, vol. 348, no. 20, 2 April 2003 (2003-04-02), pages 2030-2031, XP002288826 ISSN: 1533-4406 the whole document	1-27
A	EP 0 278 541 A (DUPHAR INT RES) 17 August 1988 (1988-08-17) the whole document	1-27
A	WO 03/013599 A (AGRONOMIQUE INST NAT RECH ; ENVT TOULOUSE (FR); GONON VALERIE (FR); AU) 20 February 2003 (2003-02-20) the whole document	1-27
X	PEI JIANWU ET AL: "Chicken interferon type I inhibits infectious bronchitis virus replication and associated respiratory illness" JOURNAL OF INTERFERON AND CYTOKINE RESEARCH, vol. 21, no. 12, December 2001 (2001-12), pages 1071-1077, XP002302419 ISSN: 1079-9907	28,29, 33,34, 36,40
Y	abstract	30-32, 35, 37-39, 41,42
X	SMITH A L ET AL: "INTRANASALLY ADMINISTERED ALPHA-BETA INTERFERON PREVENTS EXTENSION OF MOUSE HEPATITIS VIRUS STRAIN JHM INTO THE BRAINS OF BALB-CBYJ MICE" ANTIVIRAL RESEARCH, vol. 8, no. 5-6, 1987, pages 239-246, XP008037792 ISSN: 0166-3542	28,29, 33,36,40
Y	page 242	30-32, 35, 37-39, 41,42
	----- -/--	

## INTERNATIONAL SEARCH REPORT

International Application No

/NL2004/000229

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TURNER R B ET AL: "PREVENTION OF EXPERIMENTAL CORONAVIRUS COLDS WITH INTRANASAL ALPHA-2B INTERFERON" JOURNAL OF INFECTIOUS DISEASES, vol. 154, no. 3, 1986, pages 443-447, XP008037795 ISSN: 0022-1899	28,29, 31, 33-35,40
Y	figure 1  table 1	30,32, 36-39, 41,42
P,X	----- HAAGMANS BART L ET AL: "Pegylated interferon-alpha protects type 1 pneumocytes against SARS coronavirus infection in macaques." NATURE MEDICINE, vol. 10, no. 3, March 2004 (2004-03), pages 290-293, XP002302418 ISSN: 1078-8956 the whole document	28-42
P,X	----- ANONYMOUS: "Viragen files patent application for drug to target SARS" INTERNET ARTICLE, 'Online!' 2 June 2003 (2003-06-02), pages 1-2, XP002302417 Retrieved from the Internet: URL:http://www.amarbio.com/10News56.html> 'retrieved on 2004-10-25!	28,29, 33,34, 40-42
Y	the whole document	30-32, 35-39
P,X	----- STRÖHER UTE ET AL: "Severe acute respiratory syndrome-related coronavirus is inhibited by interferon- alpha." THE JOURNAL OF INFECTIOUS DISEASES. 1 APR 2004, vol. 189, no. 7, 1 April 2004 (2004-04-01), pages 1164-1167, XP008037651 ISSN: 0022-1899 the whole document	28,29, 31, 33-35, 37,40
P,Y	----- V KURIAN: "Can coconut drive SARS to nuts?" INTERNET ARTICLE, 'Online!' 20 May 2003 (2003-05-20), page 1, XP002302422 Retrieved from the Internet: URL:http://www.blonnet.com/2003/05/20/stories/2003052000241700.htm> 'retrieved on 2004-10-25! paragraph '0004!	42
	----- -/--	

# INTERNATIONAL SEARCH REPORT

ational Application No

/NL2004/000229

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with Indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>MASCI PAUL ET AL: "New and modified interferon alfas: preclinical and clinical data." CURRENT ONCOLOGY REPORTS. MAR 2003, vol. 5, no. 2, March 2003 (2003-03), pages 108-113, XP008037765 ISSN: 1523-3790 the whole document</p> <p>-----</p>	30-32

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/NL2004/000229

## Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
  
Although claims 21, 22 and 40-42 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

/NL2004/000229

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 0278541	A	17-08-1988	DK 13988 A	17-07-1988
			EP 0278541 A1	17-08-1988
			HU 45559 A2	28-07-1988
			JP 63188390 A	03-08-1988
WO 03013599	A	20-02-2003	FR 2828405 A1	14-02-2003
			EP 1429803 A2	23-06-2004
			WO 03013599 A2	20-02-2003